

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EWSPCSVTCGNGIQVRIK 18
 |||||
 FT NON TER 1
 Db 27 EWSPCSVTCGNGIQVRIK 44

RESULT 13

Q9U0P5 PRELIMINARY; PRT; 80 AA.

AC Q9U0P5: 01-MAY-2000 (TRENBLREL. 13, Created)
 DT 01-MAY-2000 (TRENBLREL. 13, Last sequence update)
 DT 01-MAR-2003 (TRENBLREL. 23, Last annotation update)
 DE Circumsporozite protein (Fragment).
 GN CS.
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5833;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=D4363;
 RA de Stricker K., Vuust J., Jepsen S., Oeuvray C., Theisen M.;
 RT "Sequence variation in the non-repeat region of the Plasmodium
 falciparum glutamate rich protein (GLURP) from Brazil, Senegalese, and
 RT Burmese field isolates and from laboratory strains."
 RL Submitted (SEP-1999) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AJ269945; CAB64184.1; -;
 DR InterPro; IPR003067; Circmsprzoite.
 DR InterPro; IPR000884; TSP1.
 DR Pfam; PF00090; TSP1, 1.
 DR PRINTS; PR01303; CRCMSPRZOITE.
 DR SMART; SM00209; TSP1, 1.
 DR PROSITE; PS50092; TSP1, 1.
 FT NON TER 1
 FT NON TER 80
 SQ SEQUENCE 80 AA; 9000 MW; 03798BD18F0BF3B3 CRC64;

Query Match 100.0%; Score 103; DB 5; Length 80;
 Best Local Similarity 100.0%; Pred. No. 8.4e-10;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EWSPCSVTCGNGIQVRIK 18
 |||||
 FT NON TER 1
 Db 27 EWSPCSVTCGNGIQVRIK 44

RESULT 14

Q9TW76 PRELIMINARY; PRT; 80 AA.

AC Q9TW76: 01-MAY-2000 (TRENBLREL. 13, Created)
 DT 01-MAY-2000 (TRENBLREL. 13, Last sequence update)
 DT 01-MAR-2003 (TRENBLREL. 23, Last annotation update)
 DE Circumsporozite protein (Fragment).
 GN CS.
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5833;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PALO ALTO, and D50;
 RA de Stricker K., Vuust J., Jepsen S., Oeuvray C., Theisen M.;
 RT "Sequence variation in the non-repeat region of the Plasmodium
 falciparum glutamate rich protein (GLURP) from Brazil, Senegalese, and
 RT Burmese field isolates and from laboratory strains."
 RL Submitted (SEP-1999) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AJ269954; CAB64179.1; -;
 DR EMBL; AJ269944; CAB64170.1; -;
 DR InterPro; IPR003067; Circmsprzoite.
 DR InterPro; IPR000884; TSP1.
 DR Pfam; PF00090; TSP1, 1.
 DR PRINTS; PR01303; CRCMSPRZOITE.

DR SMART; SM00209; TSP1, 1.
 DR PROSITE; PS50092; TSP1, 1.
 FT NON TER 1
 FT NON TER 80
 SQ SEQUENCE 80 AA; 9043 MW; 17282E319A508B3 CRC64;

Query Match 100.0%; Score 103; DB 5; Length 80;
 Best Local Similarity 100.0%; Pred. No. 8.4e-10;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EWSPCSVTCGNGIQVRIK 18
 |||||
 FT NON TER 1
 Db 27 EWSPCSVTCGNGIQVRIK 44

RESULT 15

Q9U0Q4 PRELIMINARY; PRT; 80 AA.

AC Q9U0Q4: 01-MAY-2000 (TRENBLREL. 13, Created)
 DT 01-MAY-2000 (TRENBLREL. 13, Last sequence update)
 DT 01-MAR-2003 (TRENBLREL. 23, Last annotation update)
 DE Circumsporozite protein (Fragment).
 GN CS.
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5833;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=B1;
 RA de Stricker K., Vuust J., Jepsen S., Oeuvray C., Theisen M.;
 RT "Sequence variation in the non-repeat region of the Plasmodium
 falciparum glutamate rich protein (GLURP) from Brazil, Senegalese, and
 RT Burmese field isolates and from laboratory strains."
 RL Submitted (SEP-1999) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AJ269941; CAB64167.1; -;
 DR InterPro; IPR003067; Circmsprzoite.
 DR InterPro; IPR000884; TSP1.
 DR Pfam; PF00090; TSP1, 1.
 DR PRINTS; PR01303; CRCMSPRZOITE.
 DR SMART; SM00209; TSP1, 1.
 DR PROSITE; PS50092; TSP1, 1.
 FT NON TER 1
 FT NON TER 80
 SQ SEQUENCE 80 AA; 9102 MW; A3283B70CE850FDE CRC64;

Query Match 100.0%; Score 103; DB 5; Length 80;
 Best Local Similarity 100.0%; Pred. No. 8.4e-10;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EWSPCSVTCGNGIQVRIK 18
 |||||
 FT NON TER 1
 Db 27 EWSPCSVTCGNGIQVRIK 44

Search completed: December 23, 2003, 15:35:34
 Job time : 43 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 23, 2003, 15:33:08 ; Search time 17.75 Seconds
(without alignments)
53.654 Million cell updates/sec

Title: US-09-980-564-12

Perfect score: 32
Sequence: 1 KLUKPG 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

A Geneseq_19Jun03: *
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT: *
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT: *
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT: *
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT: *
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT: *
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT: *
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT: *
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT: *
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT: *
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT: *
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT: *
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT: *
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT: *
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT: *
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT: *
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT: *
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT: *
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT: *
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT: *
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT: *
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT: *
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT: *
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT: *
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	32	100.0	6	22 AAB49236	Peptide RI derived
2	32	100.0	13	21 AAY70281	Plasmodium falciparum
3	32	100.0	14	7 AAP61001	Sequence corresp.
4	32	100.0	19	9 AAP82541	Region of P. falciparum
5	32	100.0	20	14 AAR37974	P. vivax circumsporozoite a
6	32	100.0	160	11 AAR07288	Circumsporozoite a
7	32	100.0	180	11 AAR07290	Circumsporozoite a
8	32	100.0	184	11 AAR07289	Circumsporozoite a
9	32	100.0	250	7 AAP60412	Synthetic peptide

10	32	100.0	309	12 AAR13175	NS1_81-RLfclta9.
11	32	100.0	319	11 AAR07945	NS18RLfclta9 plasm
12	32	100.0	319	12 AAR13176	NS1_81-RLfclta9. P
13	32	100.0	327	12 AAR13177	NS1_81-RLfclta9 + (
14	32	100.0	335	12 AAR13178	NS1_81 (NANP) 4RLfcl
15	32	100.0	335	12 AAR13179	NS1_81 (NANP) 4RLfcl
16	32	100.0	350	21 AAY70278	Recombinant vacci
17	32	100.0	378	8 AAP70708	Sequence encoding
18	32	100.0	378	14 AAR30609	Plasmodium vivax c
19	32	100.0	411	9 AAP83144	Sequence encoded b
20	32	100.0	412	7 AAP60416	CS protein of mala
21	32	100.0	412	9 AAP80835	Sequence encoded b
22	32	100.0	1807	22 AAB85697	Recombinant protei
23	32	100.0	2028	22 AAB85698	Recombinant protei
24	30	93.8	319	22 AAU16209	Human novel secret
25	30	93.8	319	24 ABUS5278	Human novel polype
26	30	93.8	547	22 ABG16739	Novel human diagno
27	30	93.8	949	23 ABP65113	Hypoxia-induced pr
28	29	90.6	126	22 AAU66894	Human polypeptide
29	29	90.6	130	22 AAU66894	Protonibacterium
30	29	90.6	737	15 AAR56494	TATA-binding prote
31	29	90.6	737	17 AAU66894	Human TATA-binding
32	29	90.6	737	18 AAR25019	TATA-binding prote
33	29	90.6	990	15 AAR51899	Human embryonal ki
34	29	90.6	998	24 ABU07849	Human ephrin recep
35	29	90.6	1007	22 ABB11903	Human protein Tyr
36	28	87.5	10	16 AAR66160	Murine Flt3 ligand
37	28	87.5	118	22 AAB70168	519 single chain F
38	28	87.5	123	19 AAR70619	Anti-VEGF murine a
39	28	87.5	123	23 AAB61188	Murine anti-VEGF m
40	28	87.5	185	22 AAG91568	C glutamicum prote
41	28	87.5	249	23 AAB45880	Human Blys binding
42	28	87.5	249	23 AAB45883	Human Blys binding
43	28	87.5	249	23 AAB45893	Human Blys binding
44	28	87.5	249	23 AAB46074	Human Blys binding
45	28	87.5	249	23 AAB46109	Human Blys binding

ALIGNMENTS

RESULT 1
AAB49236
ID AAB49236 standard; Protein, 6 AA.
AC AAB49236;
XX
XX
DT 14-MAR-2001 (first entry)
XX
XX
DE Peptide RI derived from malaria circumsporozoite surface protein.
XX Adenoviral vector; inverted terminal repeat; ITR; gene therapy;
XX sickle cell anemia; thalassemia; atherosclerosis; restenosis;
XX cancer.
XX
OS Synthetic.
XX
XX WO200073478-A2.
XX
XX
XX 07-DEC-2000.
XX
XX
XX 01-JUN-2000; 2000WO-US15442.
XX
XX
XX 01-JUN-1999; 99US-0137213.
XX 22-OCT-1999; 99US-0161097.
XX
XX (UNIM) UNIV WASHINGTON.
XX
XX
XX Lieber A, Shayakhmetov D, Farrar D, Papayannopoulou T;
XX WPI; 2001-049942/06.
XX
XX Recombinant adenoviral vector containing transgene, used to produce

PT gutless vectors for gene therapy, targetable to selected cells and
PT lacking antigenicity -
XX
XX Example 2; Page 97; 156pp; English.
XX
CC The present invention relates to a first generation recombinant
CC adenoviral vector, part of which integrates into a host cell genome. Two
CC inverted terminal repeats (ITRs) allow integration of a transgene into
CC the host genome. The invention can be used to make 'gutless' vectors
CC for gene therapy, e.g. of sickle cell anemia or thalassemia (targeting
CC hematopoietic cells) or atherosclerosis or restenosis (targeting
CC endothelial cells), or more generally a wide range of genetic diseases,
CC cancers and infectious diseases.
XX
SQ Sequence 6 AA;
Query Match 100.0%; Score 32; DB 22; Length 6;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KLKOPG 6
Db 1 KLKOPG 6
RESULT 2
AAY70281
ID AAY70281 standard; peptide; 13 AA.
XX
AC AAY70281;
XX
DT 06-JUN-2000 (first entry)
XX
DE Plasmodium falciparum CSP antigenic epitope, P592.
XX
KM Recombinant protein; CDC/NITMALVAC-1; multivalent; malaria; vaccine;
KM T-cell epitope; tetanus toxoid; antigenic epitope; treatment;
KM circumsporozoite protein; CSP; sporozoite surface protein-2; SSP-2;
KM liver stage antigen-1; LSA-1; merozoite surface protein-1; MSP-1; MSP-2;
KM apical membrane antigen-1; AMA-1; erythrocyte binding antigen-175;
KM EBA-175; rhoptry associated protein-1; RAP-1; gamete specific antigen;
KM Pf527; antiparasitic; prevention; anti-CDC/NITMALVAC-1 antibody.
XX
OS Plasmodium falciparum.
XX
PN WO200011179-A1.
XX
PD 02-MAR-2000.
XX
PF 19-AUG-1999; 99WO-US18869.
XX
PR 21-AUG-1998; 98US-0097703.
XX
PA (NAIM-) NAT INST IMMUNOLOGY.
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Lal AA, Shi YP, Hasnain SE;
XX
DR WPI; 2000-237654/20.
XX
PT Novel recombinant protein as vaccine for treating malarial infection
PT comprises antigenic peptides obtained from different stages of
PT plasmodium falciparum life cycle -
XX
PS Claim 2; Page 16; 52pp; English.
XX
CC The present sequence is the antigenic epitope P592, derived from
CC circumsporozoite protein (CSP) of the sporozoite stage of Plasmodium
CC falciparum. It is used in the construction of recombinant protein
CC CDC/NITMALVAC-1, which is a multivalent, multistage malarial vaccine.
CC The recombinant protein comprises, melittin signal peptide, (His) 6 tag,
CC T-cell epitope from tetanus toxoid and 21 antigenic epitopes from
CC circumsporozoite protein (CSP), sporozoite surface protein-2 (SSP-2),

CC liver stage antigen-1 (LSA-1), merozoite surface protein-1 (MSP-1),
CC MSP-2, apical membrane antigen-1 (AMA-1), erythrocyte binding antigen-175
CC (EBA-175), rhoptry associated protein-1 (RAP-1) and gamete specific
CC antigen, Pf527. These epitopes were obtained at different stages of the
CC life cycle of P. falciparum. CDC/NITMALVAC-1 vaccine has antiparasitic
CC activity and can used for treatment and prevention of malarial
CC infections. Anti-CDC/NITMALVAC-1 antibodies can be used for detecting
CC P. falciparum in biological samples.
XX
SQ Sequence 13 AA;
Query Match 100.0%; Score 32; DB 21; Length 13;
Best Local Similarity 100.0%; Pred. No. 3.4;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KLKOPG 6
Db 4 KLKOPG 9
RESULT 3
AAP61001
ID AAP61001 standard; Peptide; 14 AA.
XX
AC AAP61001;
XX
DT 25-MAR-2003 (updated)
DT 23-AUG-1991 (first entry)
XX
DE Sequence corresp. to a non-repetitive surface epitope of the
DE circumsporozoite protein (CS) of a member of the genus plasmodium.
XX
KM Malaria vaccine; reactive epitope; protective epitope.
XX
OS Plasmodium.
XX
PN WO8601721-A.
XX
PD 27-MAR-1986.
XX
PF 30-AUG-1985; 85MO-0006655.
XX
PR 12-SEP-1984; 84US-0649903.
PR 10-OCT-1985; 85CN-0107448.
PR 26-OCT-1987; 87US-0115634.
XX
PA (UINY) UNIV NEW YORK STATE.
XX
PI Vergara U, Ruiz A, Ferreira A, Nussenzwei RS, Nussenzweig VN;
XX
DR WPI; 1986-093998/14.
XX
PT Peptide for making synthetic malaria vaccine - comprising
PT cross-reactive and protective epitopes of circumsporozoite
PT proteins of plasmodium
XX
PS Claim 11; Page 18; 26pp; English.
XX
CC The peptides of the invention elicit formation of antibodies which
CC bind to the CS proteins of the malarial species from which they were
CC derived and also to the CS proteins of other malarial species. They
CC are useful in the development of a synthetic malaria vaccine and can
CC be made by synthetic methods or can form part of genetically
CC engineered constructs.
CC (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ Sequence 14 AA;
Query Match 100.0%; Score 32; DB 7; Length 14;
Best Local Similarity 100.0%; Pred. No. 3.7;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KLKOPG 6

Db 8 KLRPG 13

RESULT 4

AAP82541
ID AAP82541 standard; peptide; 19 AA.

AC AAP82541;

DT 12-DEC-1990 (first entry)

DE Region of P.falciparum circumsporozoite protein with linker attached.

KW P.falciparum circumsporozoite protein; malaria;

OS synthetic.

FN D83723583-A.

PD 28-JAN-1988.

PF 16-JUL-1987; 87DE-3723583.

PR 16-JUL-1986; 86IT-0021144.

PA (ENIE) ENRICERCH SPA.

PI Bernardi A, Bonelli F, Pessi A, Verdini AS;

DR WPI; 1988-030026/05.

XX New immunogenic polypeptide for malaria vaccine - contg.

PT sequences derived from P. falciparum circumsporozoite protein,

PT also useful for antibody assay

PS Claim 1; Page 2; 13pp; German.

XX This is the N-terminal of an immunogenic polypeptide useful as a

CC malaria vaccine. It comprises a fragment of the P.falciparum

CC circumsporozoite protein (CSP) and one copy of a peptide linker,

CC which is pref. present in 3-10 copies, which links the preceding

CC sequence to a 2nd (C-terminal) sequence comprising tetrapeptide

CC units of CSP. See AAP82542.

XX

SO Sequence 19 AA;

QY Query Match 100.0%; Score 32; DB 9; Length 19;

Best Local Similarity 100.0%; Pred. No. 5;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 KLRPG 6

6 KLRPG 11

RESULT 5

AAR37974

ID AAR37974 standard; peptide; 20 AA.

XX AAR37974;

XX 25-MAR-2003 (updated)

DT 12-OCT-1993 (first entry)

DE P. vivax circumsporozoite protein epitope peptide.

XX Human T-cell epitope; Plasmodium vivax; CS; vaccine; anti-malarial;

XX malaria; antimalarial; diagnosis; infection; past; present; monitor;

XX progress; vaccination; treatment; lymphocyte diagnostic assay;

XX Belem strain.

XX Synthetic.

XX W09311157-A1.

XX 10-JUN-1993.

XX 27-NOV-1992; 92WO-AU00639.

XX 27-NOV-1991; 91AU-0009720.

XX (COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.

XX Bilborough J, Good M;

DR WPI; 1993-196995/24.

XX Peptide for antimalarial vaccine, for monoclonal and polyclonal

PT antibodies - comprises human T cell epitope of circumsporozoite

PT protein of Plasmodium vivax or part, fragment, deriv. homologue

PT or analogue, for lymphocyte diagnostic assay

XX Example; Page 30; 50pp; English.

CC The peptide is equivalent to residues 86-105 of the circumsporozoite

CC (CS) protein of Plasmodium vivax, Belem strain. It is a T-cell epitope

CC of the CS protein and can be used in antimalarial vaccines. It can also

CC be used diagnostically, e.g. to screen for present or past P. vivax

CC infection, or to monitor the progress of vaccination or treatment.

CC The ant. required to induce T-cell immunity is pref. 10ug-1mg/kg.

CC Monoclonal and polyclonal antibodies against the peptide may be used

CC in immunotherapy and in diagnostic immunoassays.

CC (Updated on 25-MAR-2003 to correct PN field.)

XX

SO Sequence 20 AA;

QY Query Match 100.0%; Score 32; DB 14; Length 20;

Best Local Similarity 100.0%; Pred. No. 5.2;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 KLRPG 6

6 KLRPG 11

RESULT 6

AAR07288

ID AAR07288 standard; protein; 160 AA.

XX AAR07288;

XX 25-MAR-2003 (updated)

DT 28-JAN-1991 (first entry)

DE Circumsporozoite analogue Vivax 3.1.

XX CS protein; plasmodium; malaria; vaccine.

XX Synthetic.

XX EP392820-A.

XX 17-OCT-1990.

XX 11-APR-1990; 90EP-0303907.

XX 11-APR-1989; 89US-0336288.

XX (CHIR) CHIRON CORP.

XX Barr PJ, Barthurst IC, Gibson HL;

XX WPI; 1990-314486/42.

XX N-PSDB; AAQ07288.

XX

PT Recombinant plasmodium circumsporozoite analogues - lacking
PT one or more repeat epitope(s) for use as a malaria sub-unit
XX vaccine.
PS Claim 6; Fig 5; 22pp; English.
XX
CC The protein was produced by expression of a synthetic gene. The
CC analogue comprises AAs 23-91 of native P.vivax CS protein followed
CC by a single "A" sequence (DRADGQPAQ), followed by a single "B"
CC sequence (DRADGQPAQ), followed by AAs 264-335. Reduction of the
CC immunodominance of the repeat region relative to the epitopes in
CC the regions flanking the repeats enhances sporozoite neutralising
CC activity.
CC See also AAR07287-91.
CC (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ Sequence 160 AA;

QY Query Match 100.0%; Score 32; DB 11; Length 160;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 KLRQPG 6
65 KLRQPG 70

RESULT 7
AAR07290
ID AAR07290 standard; protein; 180 AA.
XX
AC AAR07290;
XX
DT 25-MAR-2003 (updated)
DT 28-JAN-1991 (first entry)
XX
DE Circumsporozoite analogue Falciparum 4.
XX
KM CS protein; plasmodium; malaria; vaccine.
XX
OS Synthetic.
XX
PN EP392820-A.
XX
PD 17-OCT-1990.
XX
PF 11-APR-1990; 90EP-0303907.
XX
PR 11-APR-1989; 89US-0336288.
XX
PA (CHIR) CHIRON CORP.
XX
PI Barr PJ, Bathurst IC, Gibson HL;
XX
XX WPI; 1990-314486/42.
DR N-PSDB; AAQ06153.
XX
XX
PT Recombinant plasmodium circumsporozoite analogues - lacking
PT one or more repeat epitope(s) for use as a malaria sub-unit
XX vaccine.
PS Claim 10; Fig 10; 22pp; English.
XX
CC The protein was produced by expression of a synthetic gene. The
CC analogue comprises AAs 68-123 of the native P. falciparum CS
CC protein, followed by four repeat sequences (three "B"s, i.e. NANP)
CC and one "A", i.e. NVDP), followed by AAs 289-392 of the native
CC protein. Reduction of the immunological dominance of the repeats
CC relative to the epitopes in the regions flanking the repeats
CC enhances sporozoite neutralising activity.
CC See also AAR07287-91.
CC (Updated on 25-MAR-2003 to correct PA field.)
XX

SQ Sequence 180 AA;

QY Query Match 100.0%; Score 32; DB 11; Length 180;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 KLRQPG 6
49 KLRQPG 54

RESULT 8
AAR07289
ID AAR07289 standard; protein; 184 AA.
XX
AC AAR07289;
XX
DT 25-MAR-2003 (updated)
DT 28-JAN-1991 (first entry)
XX
DE Circumsporozoite analogue Falciparum 3.
XX
KM CS protein; plasmodium; malaria; vaccine.
XX
OS Synthetic.
XX
PN EP392820-A.
XX
PD 17-OCT-1990.
XX
PF 11-APR-1990; 90EP-0303907.
XX
PR 11-APR-1989; 89US-0336288.
XX
PA (CHIR) CHIRON CORP.
XX
PI Barr PJ, Bathurst IC, Gibson HL;
XX
XX WPI; 1990-314486/42.
DR N-PSDB; AAQ06152.
XX
XX
PT Recombinant plasmodium circumsporozoite analogues - lacking
PT one or more repeat epitope(s) for use as a malaria sub-unit
XX vaccine.
PS Claim 10; Fig 8; 22pp; English.
XX
CC The protein was produced by expression of a synthetic gene. The
CC analogue comprises AAs 43-123 of the native P. falciparum CS
CC protein, followed by four repeat sequences (three "B"s, i.e. NANP)
CC and one "A", i.e. NVDP), followed by AAs 289-374 of the native
CC protein. Reduction of the immunological dominance of the repeats
CC relative to the epitopes in the regions flanking the repeats
CC enhances sporozoite neutralising activity.
CC See also AAR07287-91.
CC (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ Sequence 184 AA;

QY Query Match 100.0%; Score 32; DB 11; Length 184;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 KLRQPG 6
71 KLRQPG 76

RESULT 9
AAP60412
ID AAP60412 standard; protein; 250 AA.
XX
AC AAP60412;
XX

```

XX 25-MAR-2003 (updated)
DT 13-JUN-1991 (first entry)
XX
XX Synthetic peptide antigenic for the malaria parasite.
XX Sporozite; vaccination.
XX Plasmodium falciparum.
XX
XX Key Location/Qualifiers
FH Region 18..21
FT /label= Repeat unit
FT /note= "of Claim 1"
FT 22..26
FT /label= Repeat unit
FT /note= "of Claim 1"
XX
XX EPI66410-A.
XX
XX 02-JAN-1986.
XX
XX 24-JUN-1985; 85EP-0107794.
XX
XX 26-JUN-1994; 84US-0624564.
XX
XX (USDC ) US SEC OF COMMERCE.
XX (USGO ) US GOVERNMENT.
XX (USSA ) US SEC OF ARMY.
XX
XX McCutchan TF, Dame JB, Williams JL, Schneider I,
XX
XX WPI; 1986-008635/02.
XX
XX New immunologically active pure synthetic peptide(s) - used for
XX protection against infection by malaria parasite.
XX
XX Claim 7; Page 40-41; 49pp; English.
XX
XX The peptides may be synthesised in pure form and used to generate an
XX immune response in vaccination against malaria. The featured repeat
XX units are claimed and must be present in copies of 2-1000, in the
XX P.falciparum genome, the first is coded for 37 times.
XX See also AAN60362.
XX (Updated on 25-MAR-2003 to correct PA field.)
XX
XX Sequence 250 AA;
SQ
Query Match 100.0%; Score 32; DB 7; Length 250;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 KTKOPG 6
Db 6 KTKOPG 11

```

```

XX Key Location/Qualifiers
FH Region 1..81
FT /label= N-terminal of NS1
FT /note= "Influenza virus nonstructural protein 1"
FT
FT Peptide 82..87
FT /label= synthetic linker
FT 88
FT /label= artifact
FT /note= "see comments"
FT
FT Region 89..193
FT /label= Aas 19-123 of CS protein
FT /note= "Region I contg. flanking region less
FT signal sequence"
FT 194..309
FT /label= Aas 297-412 of CS protein
FT /note= "Region II flanking region minus 9 N-term-
FT inal Aas"
XX
XX EPI32965-A.
XX
XX 19-JUN-1991.
XX
XX 06-DEC-1990; 90EP-0313257.
XX
XX 08-DEC-1989; 89US-0447746.
XX
XX (SMITK ) SWITKLINE BECHAM.
XX (USSA ) US SEC OF ARMY.
XX (BIOM-) BIOMEDICAL RES INST.
XX (GROS/) GROSS M S.
XX
XX Gross MS, Gordon DM, Hollingdale MR;
XX
XX WPI; 1991-179771/25.
XX
XX Polypeptide comprising immunogenic determinants from P falciparum
XX - for vaccine against malaria infection in humans.
XX
XX Example 1; Page 7; 18pp; English.
XX
XX The polypeptide is prepd. by genetic engineering of genes encoding
XX the P. falciparum circumsporozoite (CS) protein [Dame et al.,
XX Science 225 : 593 (1984)], and the influenza virus non-structural
XX protein 1 (NS1) [Baez et al., Nucleic Acids Research, 8 : 5845
XX (1980)]. The DNA encoding the 1st 81 AAs of the N-terminal of NS1
XX (NS1_81) is linked via a synthetic sequence to DNA encoding Region
XX I contg. flanking regionless the 18 AA signal region, which in
XX turn is fused to DNA encoding Region II contg. flanking region
XX less the first nine N-terminal AAs. This CS fusion is designated
XX RIdelta9. The Pro residue separating the Asp (at the C-terminal
XX of the linker) from RIdelta9 is an artifact of a filled in BamHI
XX site. The peptide can be used in a vaccine for protection against
XX malaria.
XX See also AAR12306-R12311 and AAR13176-R13179.
XX (Updated on 25-MAR-2003 to correct PA field.)
XX
XX Sequence 309 AA;
SQ
Query Match 100.0%; Score 32; DB 12; Length 309;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 KTKOPG 6
Db 182 KTKOPG 187

```

```

RESULT 11
AAR07945
ID AAR07945 standard; protein; 319 AA.
XX
XX AAR07945;

```

```

XX 25-MAR-2003 (updated)
DT 22-FEB-1991 (first entry)
XX
DE NS181RLFauth plasmid product.
XX
KM Malaria; vaccine.
XX
OS Plasmodium falciparum.
XX
FH Key Location/Qualifiers
FT Domain 1..81
FT /label= NS181 protein fragment
FT /note= "from plasmid pMG-1"
FT Domain 89..193
FT /label= Fragment of circumsporozite protein
FT Domain 204..319
FT /label= Fragment of circumsporozite protein
XX
EP398540-A.
XX
PD 22-NOV-1990.
XX
PF 01-MAY-1990; 90EP-0304720.
XX
PR 03-MAY-1989; 89US-0346863.
XX
PA (SMIK ) SMITHKLINE BEECHAM.
PA (GROS/) GROSS M S.
XX
PI Grosse MS, Young JF;
XX
DR WPI; 1990-350299/47.
XX
DR N-PSDB; AAQ06580.
XX
PT New polypeptide used in malaria vaccine - comprises immunogenic
PT determinant from 1st and 2nd flanking regions of Plasmodium
PT surface protein and intermediate repeat domain
XX
PS Example 2; Page 11-12; 24pp; English.
XX
CC The product is useful in preparation of vaccines for treatment and
CC prophylaxis of Plasmodium sporozite infection. It may be easily
CC produced in large pure quantities from a transformed E.coli
CC expression system.
CC (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ Sequence 319 AA;
XX
Query Match 100.0%; Score 32; DB 11; Length 319;
Best Local Similarity 100.0%; Pred. No. 81;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 KTKOPG 6
Db 182 KTKOPG 187

```

RESULT 12
ID AAR13176 standard; Protein; 319 AA.

XX AAR13176;
XX
XX 25-MAR-2003 (updated)
DT 29-AUG-1991 (first entry)
XX
DE NS1_81-RLFaith.
XX
KM Immunogenic determinant; circumsporozite; CS; vaccine; malaria;
KM hybrid; influenza virus; non-structural protein 1; fusion.
XX
OS Plasmodium falciparum.

```

OS Influenza virus (A/PR/8/34) .
XX
XX Key Location/Qualifiers
XX Region 1..81
XX /label= N-terminal of NS1
XX /note= "Influenza virus nonstructural protein 1"
XX
FT Peptide 82..87
FT /label= synthetic linker
XX
FT Region 88
FT /label= artifact
FT /note= "see comments"
XX
FT Region 89..193
FT /label= AAs 19-123 of CS protein
FT /note= "Region 1 contg. flanking region less
FT signal sequence"
XX
FT Region 194
FT /label= artifact
FT /note= "see comments"
XX
FT Region 195..319
FT /label= AAs 288-412 of CS protein
FT /note= "Region II flanking region"
XX
EP432965-A.
XX
PD 19-JUN-1991.
XX
PF 06-DEC-1990; 90EP-0313257.
XX
PR 08-DEC-1989; 89US-0447746.
XX
PA (SMIK ) SMITHKLINE BEECHAM.
PA (USDA ) US SEC OF ARMY.
PA (BIOM-) BIOMEDICAL RES INST.
PA (GROS/) GROSS M S.
XX
XX
PI Grosse MS, Gordon DM, Hollingdale MR;
XX
DR WPI; 1991-179771/25.
XX
PT Polypeptide comprising immunogenic determinants from P falciparum
PT - for vaccine against malaria infection in humans.
XX
PS Example 2; Page 10; 16pp; English.
XX
XX
CC The polypeptide is prepd. by genetic engineering of genes encoding
CC the P. falciparum circumsporozite (CS) protein [Dane et al.,
CC Science 225 : 593 (1984)], and the influenza virus non-structural
CC protein 1 (NS1) [Baez et al., Nucleic Acids Research, 8 : 5845
CC (1980)]. The DNA encoding the 1st 81 AAs of the N-terminal of NS1
CC (NS1_81) is linked via a synthetic sequence to DNA encoding Region
CC I contg. flanking regionless the 18 AA signal region, which in
CC turn is fused to DNA encoding Region II-contg. flanking region.
CC This CS fusion is designated RLFAuth. The Pro residue separating
CC the Asp (at the C-terminal of the linker) from RLFAuth is an arti-
CC fact of a filled in BamHI site; the Gly separating Region I and
CC Region II-contg. CS flanking regions is an artifact of a synthetic
CC FokI/NotI linker. The peptide can be used in a vaccine for
CC protection against malaria.
CC The complete nucleotide and AA sequences are given in EP-304720,
CC filed May 1, 1990.
CC See also AAR12306-R12311 and AAR13175-R13179.
CC (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ Sequence 319 AA;
XX
Query Match 100.0%; Score 32; DB 12; Length 319;
Best Local Similarity 100.0%; Pred. No. 81;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 KTKOPG 6
Db 182 KTKOPG 187

```



```

RESULT 13
AAR13177
ID AAR13177 standard; Protein; 327 AA.
XX
AC AAR13177;
XX
DT 25-MAR-2003 (updated)
DT 29-AUG-1991 (first entry)
XX
DE NS1_81-RLfauth + (NANP)2.
XX
KM Immunogenic determinant; circumsporozoite; CS; vaccine; malaria;
KM hybrid; influenza virus; non-structural protein 1; fusion.
XX
OS Plasmodium falciparum.
OS Influenza virus (A/PR/8/34/).
XX
FH Key
FH Region
FT 1..81
FT /label= N-terminal of NS1
FT /note= "Influenza virus nonstructural protein 1"
FT 82..87
FT /label= synthetic linker
FT 88
FT /label= artifact
FT /note= "see comments"
FT 89..193
FT /label= AAs 19-123 of CS protein
FT /note= "Region I contg. flanking region less
FT /note= "signal sequence"
FT 194..201
FT /label= immunodominant repeat region
FT /note= "two tetrapeptide repeat units"
FT 202
FT /label= artifact
FT /note= "see comments"
FT 203..327
FT /label= AAs 288-412 of CS protein
FT /note= "Region II flanking region"
XX
PN EP432965-A.
XX
PD 19-JUN-1991.
XX
PF 06-DEC-1990; 90EP-0313257.
XX
PR 08-DEC-1989; 89US-0447746.
XX
PA (SMIK ) SMITHKLINE BEECHAM.
PA (USSA ) US SEC OF ARMY.
PA (BIOM-) BIOMEDICAL RES INST.
PA (GROS/) GROSS M S.
XX
PI Grose MS, Gordon DM, Hollingdale MR;
XX
DR WPI, 1991-179771/25.
XX
PT Polypeptide comprising immunogenic determinants from P falciparum
PT - for vaccine against malaria infection in humans.
XX
PS Example 3; Page 10; 18pp; English.
XX
CC The polypeptide is prepd. by genetic engineering of genes encoding
CC the P. falciparum circumsporozoite (CS) protein [Dame et al.,
CC Science 225 : 593 (1984)], and the influenza virus non-structural
CC protein 1 (NS1), [Baez et al., Nucleic Acids Research, 8 : 5845
CC (1980)]. The DNA encoding the 1st 81 AAs of the N-terminal of NS1
CC (NS1_81) is linked via a synthetic sequence to DNA encoding Region
CC I contg. flanking region less the 18 AA signal region. This is
CC linked to a synthetic sequence encoding two repeat units from the
CC immunodominant region, which in turn is fused to DNA encoding the
CC Region II-contg. flanking region. The Pro residue separating the

```

```

CC Asp (at the C-terminal of the linker) from the Region I-contg. CS
CC flanking region is an artifact of a filled-in BamHI site; the Gly
CC separating the repeat units and the Region II-contg. CS flanking
CC region is an artifact of a synthetic FokI/NotI linker. The
CC peptide can be used in a vaccine for protection against malaria.
CC See also AAR13306-R12311 and AAR13175-R13179.
CC (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ Sequence 327 AA;
XX
Query Match 100.0%; Score 32; DB 12; Length 327;
Best Local Similarity 100.0%; Pred. No. 83;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 KTKQPG 6
DB 182 KTKQPG 187
XX
RESULT 14
AAR13178
ID AAR13178 standard; Protein; 335 AA.
XX
AC AAR13178;
XX
DT 25-MAR-2003 (updated)
DT 29-AUG-1991 (first entry)
XX
DE NS1_81(NANP)4RLfauth.
XX
KM Immunogenic determinant; circumsporozoite; CS; vaccine; malaria;
KM hybrid; influenza virus; non-structural protein 1; fusion.
XX
OS Plasmodium falciparum.
OS Influenza virus (A/PR/8/34).
XX
FH Key
FH Region
FT 1..81
FT /label= N-terminal of NS1
FT /note= "Influenza virus nonstructural protein 1"
FT 82..97
FT /label= immunodominant repeat region
FT /note= "four tetrapeptide repeat units"
FT 98..103
FT /label= synthetic linker
FT 104
FT /label= artifact
FT /note= "see comments"
FT 105..209
FT /label= AAs 19-123 of CS protein
FT /note= "Region I contg. flanking region less
FT /note= "signal sequence"
FT 210
FT /label= artifact
FT /note= "see comments"
FT 211..335
FT /label= AAs 288-412 of CS protein
FT /note= "Region II flanking region"
XX
PN EP432965-A.
XX
PD 19-JUN-1991.
XX
PF 06-DEC-1990; 90EP-0313257.
XX
PR 08-DEC-1989; 89US-0447746.
XX
PA (SMIK ) SMITHKLINE BEECHAM.
PA (USSA ) US SEC OF ARMY.
PA (BIOM-) BIOMEDICAL RES INST.
PA (GROS/) GROSS M S.
XX
PI Grose MS, Gordon DM, Hollingdale MR;

```

```

XX WPI; 1991-179771/25.
DR
XX Polypeptide comprising immunogenic determinants from P falciparum
PT - for vaccine against malaria infection in humans.
XX
XX Example 4; Page 11; 18pp; English.
XX
CC The polypeptide is prepd. by genetic engineering of genes encoding
CC the P. falciparum circumsporozoite (CS) protein [Dane et al.,
CC Science 225 : 593 (1984)], and the influenza virus non-structural
CC protein 1 (NS1), [Baez et al., Nucleic Acids Research, 8 : 5845
CC (1980)]. The DNA encoding the 1st 81 AAs of the N-terminal of NS1
CC (NS1_81) is linked to a synthetic sequence encoding four repeat
CC units from the immunodominant region, which in turn is linked via
CC a synthetic sequence to DNA encoding Region I contg. flanking
CC region less the 18 AA signal region. This is linked to DNA
CC encoding Region II-contg. flanking region. The Pro residue sep-
CC arating the Asp (at the C-terminal of the linker) from the Region
CC I-contg. CS flanking region is an artifact of a filled-in BamHI
CC site; the Gly separating the Region I and II-contg. CS flanking
CC regions is an artifact of a synthetic FokI/TthIII I linker. The
CC peptide can be used in a vaccine for protection against malaria.
CC See also AAR12306-R12311 and AAR13175-R13179.
CC (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ Sequence 335 AA;
XX
Query Match 100.0%; Score 32; DB 12; Length 335;
Best Local Similarity 100.0%; Pred. No. 85;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KTKQPG 6
Db 198 KTKQPG 203

```

```

FT Region 211..335
FT /label= AAs 288-412 of CS protein
FT /note= "Region II flanking region"
XX
XX EP432965-A.
XX
XX 19-JUN-1991.
XX
XX 06-DEC-1990; 90EP-0313257.
XX
XX 08-DEC-1989; 89US-0447746.
XX
XX (SMIK ) SMITHKLINE BEECHAM.
XX (USSA ) US SEC OF ARMY.
XX (BIOM-) BIOMEDICAL RES INST.
XX (GROS/) GROSS M S.
XX
XX Gross MS, Gordon DM, Hollingdale MR;
XX WPI; 1991-179771/25.
XX
XX Polypeptide comprising immunogenic determinants from P falciparum
XX - for vaccine against malaria infection in humans.
XX
XX Example 5; Page 11; 18pp; English.
XX
CC The polypeptide is prepd. by genetic engineering of genes encoding
CC the P. falciparum circumsporozoite (CS) protein [Dane et al.,
CC Science 225 : 593 (1984)], and the influenza virus non-structural
CC protein 1 (NS1), [Baez et al., Nucleic Acids Research, 8 : 5845
CC (1980)]. The DNA encoding the 1st 81 AAs of the N-terminal of NS1
CC (NS1_81) is linked to a synthetic sequence encoding four repeat
CC units (the variant form) from the immunodominant region, which in
CC turn is linked via a synthetic sequence to DNA encoding Region I
CC contg. flanking region less the 18 AA signal region. This is
CC linked to DNA encoding Region II-contg. flanking region. The Pro
CC residue separating the Asp (at the C-terminal of the linker) from
CC the Region I-contg. CS flanking region is an artifact of a filled-
CC in BamHI site; the Gly separating the Region I and II-contg. CS
CC flanking regions is an artifact of a synthetic FokI/TthIII I
CC linker. The peptide can be used in a vaccine for protection
CC against malaria.
CC See also AAR12306-R12311 and AAR13175-R13178.
CC (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ Sequence 335 AA;
XX
Query Match 100.0%; Score 32; DB 12; Length 335;
Best Local Similarity 100.0%; Pred. No. 85;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KTKQPG 6
Db 198 KTKQPG 203

```

Search completed: December 23, 2003, 15:36:57
 Job time : 18.75 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Comugen Ltd.

OM protein - protein search, using SW model

Run on: December 23, 2003, 15:33:08 ; Search time 7.25 Seconds
(without alignments)
35.016 Million cell updates/sec

Title: US-09-980-564-12

Perfect score: 32

Sequence: 1 KLKOPG 6

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents_AA:*
1: /cgn2_6/ptodata/1/aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/aa/PCTUS.COMB.pep:*
6: /cgn2_6/ptodata/1/aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	32	100.0	13	6	5178861-17
2	32	100.0	15	6	5178861-11
3	32	100.0	22	6	5178861-5
4	32	100.0	232	6	5178861-7
5	32	100.0	378	6	5178861-9
6	32	100.0	412	1	US-08-313-288B-18
7	29	90.6	737	1	US-08-188-582-16
8	29	90.6	737	1	US-08-646-715-16
9	29	90.6	998	2	US-08-449-645A-20
10	29	90.6	998	2	US-08-702-367A-20
11	29	90.6	998	5	PCT-US95-04681-20
12	28	87.5	10	5	PCT-US94-05150-4
13	28	87.5	580	4	US-09-252-991A-22053
14	27	84.4	158	4	US-09-134-001C-4336
15	27	84.4	190	1	US-07-956-700B-38
16	27	84.4	190	1	US-08-476-537-38
17	27	84.4	190	1	US-08-485-607-38
18	27	84.4	190	2	US-08-475-879-38
19	27	84.4	190	4	US-09-433-043B-38
20	27	84.4	216	4	US-09-198-452A-717
21	27	84.4	333	4	US-09-328-352-6253
22	27	84.4	376	1	US-08-253-155A-33
23	27	84.4	417	4	US-08-887-534A-38
24	27	84.4	417	4	US-08-887-534A-40
25	27	84.4	417	4	US-08-887-534A-42
26	27	84.4	417	4	US-09-527-431-38
27	27	84.4	417	4	US-09-527-431-40

28	27	84.4	417	4	US-09-527-431-42	Sequence 42, Appl
29	27	84.4	802	4	US-09-433-043B-120	Sequence 120, App
30	27	84.4	1061	4	US-09-328-352-4445	Sequence 4445, App
31	27	84.4	2237	1	US-08-354-973-1	Sequence 1, Appl
32	26	81.2	5	6	5178861-12	Patent No. 5178861
33	26	81.2	9	2	US-08-934-222-70	Sequence 70, Appl
34	26	81.2	9	2	US-08-933-402-70	Sequence 70, Appl
35	26	81.2	9	2	US-09-207-621-70	Sequence 70, Appl
36	26	81.2	9	2	US-08-532-818-70	Sequence 70, Appl
37	26	81.2	9	2	US-09-231-797-70	Sequence 70, Appl
38	26	81.2	9	3	US-08-934-224-70	Sequence 70, Appl
39	26	81.2	9	3	US-08-933-843-70	Sequence 70, Appl
40	26	81.2	9	3	US-08-934-223-70	Sequence 70, Appl
41	26	81.2	9	3	US-09-413-492-70	Sequence 70, Appl
42	26	81.2	14	6	5178861-2	Patent No. 5178861
43	26	81.2	14	6	5178861-11	Patent No. 5178861
44	26	81.2	15	1	US-08-097-997A-2	Sequence 2, Appl
45	26	81.2	15	3	US-08-665-574C-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
5178861-17
Patent No. 5178861
APPLICANT: VERGARA, UNISES, RUIZ, ANDRES, FERREIRA, ARTURO;
NUSSENZWEIG, RUTH S.; NUSSENZWEIG, VICTOR N.
TITLE OF INVENTION: CROSS-REACTIVE AND PROTECTIVE EPITOPES
OF CIRCUMPOROITE PROTEINS
NUMBER OF SEQUENCES: 18
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/370,241
FILING DATE: 22-JUN-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 115,634
FILING DATE: 26-OCT-1987
APPLICATION NUMBER: 649,903
FILING DATE: 12-SEP-1984
SEQ ID NO:17:
LENGTH: 13
5178861-17

Query Match 100.0%; Score 32; DB 6; Length 13;
Best Local Similarity 100.0%; Pred. No. 1.8; 0; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0;

QY 1 KLKOPG 6
Db 7 KLKOPG 12

RESULT 2
517843-11
Patent No. 517843
APPLICANT: NUSSENZWEIG, VICTOR
TITLE OF INVENTION: IMMUNOGENIC POLYPEPTIDE AND METHOD FOR
PURIFYING IT
NUMBER OF SEQUENCES: 13
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/175,112
FILING DATE: 30-MAR-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 754,645
FILING DATE: 9-JUL-1985
APPLICATION NUMBER: 115,634
FILING DATE: 26-OCT-1987
APPLICATION NUMBER: 649,903
FILING DATE: 12-SEP-1984
SEQ ID NO:11:
LENGTH: 15
517843-11

Query Match 100.0%; Score 32; DB 6; Length 15;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKKOPG 6
Db 10 KKKOPG 15

RESULT 3
5178861-5
; Patent No. 5178861
; APPLICANT: VERGARA, ULISES; RUIZ, ANDRES; FERREIRA, ARTURO;
; NUSSENZWEIG, RUTH S.; NUSSENZWEIG, VICTOR N.
; TITLE OF INVENTION: CROSS-REACTIVE AND PROTECTIVE EPITOPES
; OF CIRCUMPOROZOITE PROTEINS
; NUMBER OF SEQUENCES: 18
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/370,241
; FILING DATE: 22-JUN-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 115,634
; FILING DATE: 26-OCT-1987
; APPLICATION NUMBER: 649,903
; FILING DATE: 12-SEP-1984
; SEQ ID NO: 5:
; LENGTH: 22

5178861-5

Query Match 100.0%; Score 32; DB 6; Length 22;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKKOPG 6
Db 8 KKKOPG 13

RESULT 4
517843-7
; Patent No. 5171843
; APPLICANT: NUSSENZWEIG, VICTOR
; TITLE OF INVENTION: IMMUNOGENIC POLYPEPTIDE AND METHOD FOR
; PURIFYING IT
; NUMBER OF SEQUENCES: 13
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/175,112
; FILING DATE: 30-MAR-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 754,645
; FILING DATE: 9-JUL-1985
; APPLICATION NUMBER: 115,634
; FILING DATE: 26-OCT-1987
; APPLICATION NUMBER: 649,903
; FILING DATE: 12-SEP-1984
; SEQ ID NO: 7:
; LENGTH: 232

5171843-7

Query Match 100.0%; Score 32; DB 6; Length 232;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKKOPG 6
Db 10 KKKOPG 15

RESULT 5
5171843-9
; Patent No. 5171843
; APPLICANT: NUSSENZWEIG, VICTOR
; TITLE OF INVENTION: IMMUNOGENIC POLYPEPTIDE AND METHOD FOR

; PURIFYING IT
; NUMBER OF SEQUENCES: 13
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/175,112
; FILING DATE: 30-MAR-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 754,645
; FILING DATE: 9-JUL-1985
; APPLICATION NUMBER: 115,634
; FILING DATE: 26-OCT-1987
; APPLICATION NUMBER: 649,903
; FILING DATE: 12-SEP-1984
; SEQ ID NO: 9:
; LENGTH: 378

5171843-9

Query Match 100.0%; Score 32; DB 6; Length 378;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKKOPG 6
Db 92 KKKOPG 97

RESULT 6
US-08-313-288B-18
; Sequence 18, Application US/08313288B
; Patent No. 5750502
; GENERAL INFORMATION:
; APPLICANT: Jessell, Thomas M. and Avihu Klar
; TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A
; TITLE OF INVENTION: NOVEL SECRETED PROTEIN, F-SPONDIN
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/313,288B
; FILING DATE: January 5, 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 40028-A-PCT-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0526
; TELEX:
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 412 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-313-288B-18

US-08-313-288B-18

Query Match 100.0%; Score 32; DB 1; Length 412;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKKOPG 6
Db 1 KKKOPG 6

Db 112 KIKOPG 117

RESULT 7

US-08-188-582-16
Sequence 16, Application US/08188582

Patent No. 5534410

GENERAL INFORMATION:

APPLICANT: Tjian, Robert

APPLICANT: Comai, Lucio

APPLICANT: Dynlacht, Brian D.

APPLICANT: Hoey, Timothy

APPLICANT: Rupprecht, Siegfried

APPLICANT: Tanese, Naoko

APPLICANT: Weinzierl, Robert O.J.

TITLE OF INVENTION: TATA-BINDING PROTEIN ASSOCIATED FACTORS.

TITLE OF INVENTION: NUCLEIC ACIDS ENCODING TAFs AND METHODS OF USE

NUMBER OF SEQUENCES: 36

CORRESPONDENCE ADDRESS:

ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT

STREET: 4 Embarcadero Center, Suite 3400

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94111-4187

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/188,582

FILING DATE: 28-JAN-1994

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Osman, Richard A

REGISTRATION NUMBER: 36,627

REFERENCE/DOCKET NUMBER: A-57650-2/AJT/RAO

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 781-1989

TELEFAX: (415) 398-3249

TELEX: 910 277299

INFORMATION FOR SEQ ID NO: 16:

SEQUENCE CHARACTERISTICS:

LENGTH: 737 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-188-582-16

Query Match

Best Local Similarity 83.3%; Score 29; DB 1; Length 737;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 471 KIKOPG 476

RESULT 8

US-08-646-715-16
Sequence 16, Application US/08646715

Patent No. 5637686

GENERAL INFORMATION:

APPLICANT: Tjian, Robert

APPLICANT: Comai, Lucio

APPLICANT: Dynlacht, Brian D.

APPLICANT: Hoey, Timothy

APPLICANT: Rupprecht, Siegfried

APPLICANT: Tanese, Naoko

APPLICANT: Weinzierl, Robert O.J.

TITLE OF INVENTION: TATA-BINDING PROTEIN ASSOCIATED FACTORS,
TITLE OF INVENTION: NUCLEIC ACIDS ENCODING TAFs AND METHODS OF USE
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:

ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT

STREET: 4 Embarcadero Center, Suite 3400

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94111-4187

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/646,715

FILING DATE: 09-MAY-1996

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/188,582

FILING DATE: 28-JAN-1994

ATTORNEY/AGENT INFORMATION:

NAME: Osman, Richard A

REGISTRATION NUMBER: 36,627

REFERENCE/DOCKET NUMBER: A-57650-2/AJT/RAO

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 781-1989

TELEFAX: (415) 398-3249

TELEX: 910 277299

INFORMATION FOR SEQ ID NO: 16:

SEQUENCE CHARACTERISTICS:

LENGTH: 737 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-646-715-16

Query Match

Best Local Similarity 83.3%; Score 29; DB 1; Length 737;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 471 KIKOPG 476

RESULT 9

US-08-449-645A-20
Sequence 20, Application US/08449645A

Patent No. 5981245

GENERAL INFORMATION:

APPLICANT: Fox, Gary M.

TITLE OF INVENTION: Eph-like Receptor Protein Tyrosine

TITLE OF INVENTION: Kinases

NUMBER OF SEQUENCES: 43

CORRESPONDENCE ADDRESS:

ADDRESSEE: Amgen Patent Operations/RBW

STREET: 1840 Denavilland Drive

CITY: Thousand Oaks

STATE: California

COUNTRY: USA

ZIP: 91320

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/449,645A

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Winter, Robert B.
REFERENCE/DOCKET NUMBER: A-287
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 998 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-449-645A-20

Query Match 90.6%; Score 29; DB 2; Length 998;
Best Local Similarity 83.3%; Pred. No. 4.8e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 KLKOPG 6
DB 651 RLKOPG 656

RESULT 10
US-08-702-367A-20
Sequence 20, Application US/08702367A
Patent No. 5981246
GENERAL INFORMATION:
APPLICANT: Fox, Gary M.
TITLE OF INVENTION: Eph-Like Receptor Protein Tyrosine
TITLE OF INVENTION: Kinases
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amgen Patent Operations/RBW
STREET: 1840 Dehavilland Drive
CITY: Thousand Oaks
STATE: California
COUNTRY: USA
ZIP: 91320
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/702.367A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Winter, Robert B.
REFERENCE/DOCKET NUMBER: A-287
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 998 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-702-367A-20

Query Match 90.6%; Score 29; DB 2; Length 998;
Best Local Similarity 83.3%; Pred. No. 4.8e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 KLKOPG 6
DB 651 RLKOPG 656

RESULT 11
PCT-US95-04681-20
Sequence 20, Application PC/TUS9504681
GENERAL INFORMATION:
APPLICANT: Fox, Gary M.
TITLE OF INVENTION: Eph-Like Receptor Protein Tyrosine
TITLE OF INVENTION: Kinases

NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amgen Patent Operations/RBW
STREET: 1840 Dehavilland Drive
CITY: Thousand Oaks
STATE: California
COUNTRY: USA
ZIP: 91320
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/04681
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Winter, Robert B.
REFERENCE/DOCKET NUMBER: A-287
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 998 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-04681-20

Query Match 90.6%; Score 29; DB 5; Length 998;
Best Local Similarity 83.3%; Pred. No. 4.8e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 KLKOPG 6
DB 651 RLKOPG 656

RESULT 12
PCT-US94-05150-4
Sequence 4, Application PC/TUS9405150
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: Purified Mammalian Flt3 Ligands and Agonists and Antagonists
NUMBER OF SEQUENCES: 37
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh 6.0.5
SOFTWARE: Microsoft Word 5.1a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/05150
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/162,413
FILING DATE: 03-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/155,111
FILING DATE: 19-NOV-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/112,391
FILING DATE: 24-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/106,340
FILING DATE: 13-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/092,549
FILING DATE: 16-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/089,263
FILING DATE: 07-JUL-1993
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/065,231
FILING DATE: 19-MAY-1993
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
PCT-US94-05150-4

Query Match 87.5%; Score 28; DB 5; Length 10;
Best Local Similarity 83.3%; Pred. No. 9.4;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLKQPG 6
|||
DB 4 QLKQPG 9

RESULT 13
US-09-252-991A-22053
Sequence 22053, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 22053
LENGTH: 580
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22053

Query Match 87.5%; Score 28; DB 4; Length 580;
Best Local Similarity 83.3%; Pred. No. 4.6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLKQPG 6
|||
DB 7 KLKQPG 12

RESULT 14
US-09-134-001C-4336
Sequence 4336, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 4336
LENGTH: 158
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4336

Query Match 84.4%; Score 27; DB 4; Length 158;

Best Local Similarity 83.3%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 KLKQPG 6
|||
DB 110 KCKQPG 115

RESULT 15
US-07-956-700B-38
Sequence 38, Application US/07956700B
Patent No. 5539092
GENERAL INFORMATION:
APPLICANT: Robert Haselkorn and Piotr Gornicki
TITLE OF INVENTION: Cyanobacterial and Plant Acetyl-CoA
TITLE OF INVENTION: Carboxylase
NUMBER OF SEQUENCES: 116
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: 321 No. 5539092th Clark Street
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60610
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy Disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII-DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/956,700B
FILING DATE: 19921002
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Thomas E. No. 5539092thrup
REGISTRATION NUMBER: 33,268
REFERENCE/DOCKET NUMBER: ARCD:058
TELECOMMUNICATION INFORMATION:
TELEPHONE: 1-312-744-0090
TELEFAX: 1-312-755-4489
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 190 amino acids
TYPE: Amino acid
STRANDNESS: Single
TOPOLOGY: Linear
MOLECULE TYPE: Peptide
US-07-956-700B-38

Query Match 84.4%; Score 27; DB 1; Length 190;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LKQPG 6
|||
DB 139 LKQPG 143

Search completed: December 23, 2003, 15:37:38
Job time: 7.25 secs

This Page Blank (uspro)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 23, 2003, 15:33:48 ; Search time 12.25 Seconds

(without alignments)
91.474 Million cell updates/sec

Title: US-09-980-564-12

Perfect score: 32

Sequence: 1 KLRKPG 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 696363 seqs, 186758610 residues

Total number of hits satisfying chosen parameters: 696363

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/1/pubppaa/PCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/1/pubppaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/1/pubppaa/PCTUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/1/pubppaa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/1/pubppaa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/ptodata/1/pubppaa/US10A_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/1/pubppaa/US10B_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/1/pubppaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/1/pubppaa/US10_NEW_PUB.pep:*
- 17: /cgn2_6/ptodata/1/pubppaa/US60_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	32	100.0	1056	12	US-10-098-871-65
2	30	93.8	319	10	US-09-764-864-1162
3	29	90.6	614	12	US-10-054-935-2
4	29	90.6	614	12	US-10-144-194A-66
5	29	90.6	615	12	US-10-054-935-4
6	29	90.6	998	12	US-10-354-358-4
7	28	87.5	123	9	US-09-056-1608-9
8	28	87.5	123	12	US-10-234-671-9
9	28	87.5	185	10	US-09-738-626-5322
10	28	87.5	249	11	US-09-880-748-1891
11	28	87.5	249	11	US-09-880-748-1894
12	28	87.5	249	11	US-09-880-748-1904
13	28	87.5	249	11	US-09-880-748-2085
14	28	87.5	249	11	US-09-880-748-2120
15	28	87.5	249	11	US-09-880-748-2122

16	28	87.5	450	11	US-09-927-827-67	Sequence 67, Appl
17	28	87.5	546	15	US-10-173-539-10	Sequence 10, Appl
18	28	87.5	550	12	US-10-312-187-35	Sequence 30, Appl
19	28	87.5	590	12	US-09-800-187-14	Sequence 14, Appl
20	28	87.5	590	12	US-10-172-094-99	Sequence 99, Appl
21	28	87.5	590	15	US-10-072-094-99	Sequence 99, Appl
22	28	87.5	590	15	US-10-173-539-11	Sequence 11, Appl
23	28	87.5	780	12	US-10-172-094-93	Sequence 93, Appl
24	28	87.5	780	15	US-10-072-094-93	Sequence 93, Appl
25	28	87.5	835	15	US-10-172-094-90	Sequence 90, Appl
26	28	87.5	879	12	US-10-172-094-90	Sequence 90, Appl
27	28	87.5	879	15	US-10-072-094-90	Sequence 90, Appl
28	28	87.5	879	15	US-10-173-539-4	Sequence 4, Appl
29	28	87.5	879	15	US-10-173-539-6	Sequence 6, Appl
30	28	87.5	1011	12	US-10-354-358-22	Sequence 22, Appl
31	28	87.5	1011	12	US-10-172-094-89	Sequence 89, Appl
32	28	87.5	1011	15	US-10-072-094-89	Sequence 89, Appl
33	28	87.5	1011	15	US-10-173-539-2	Sequence 2, Appl
34	28	87.5	1069	12	US-10-172-094-87	Sequence 87, Appl
35	28	87.5	1069	15	US-10-072-094-87	Sequence 87, Appl
36	28	87.5	3712	14	US-10-108-605-103	Sequence 103, App
37	27	84.4	87	9	US-09-864-761-43429	Sequence 43429, A
38	27	84.4	120	12	US-10-332-122-4	Sequence 4, Appl
39	27	84.4	165	10	US-09-910-151-4	Sequence 4, Appl
40	27	84.4	167	9	US-09-925-301-1375	Sequence 1375, Ap
41	27	84.4	227	12	US-10-170-789-54	Sequence 54, Appl
42	27	84.4	227	14	US-10-045-367A-4	Sequence 4, Appl
43	27	84.4	228	12	US-10-238-075-1003	Sequence 1003, Ap
44	27	84.4	249	10	US-09-961-721-5	Sequence 5, Appl
45	27	84.4	249	12	US-10-170-789-46	Sequence 46, Appl

ALIGNMENTS

RESULT 1
US-10-098-871-65
; Sequence 65, Application US/10098871
; Publication No. US20030198958A1
GENERAL INFORMATION:
; APPLICANT: Shinketsu, Richard A.
; APPLICANT: Fernandes, Elma
; APPLICANT: Herrmann, John
; APPLICANT: Liu, Xiaohong
; APPLICANT: Yang, Meijia
; APPLICANT: Boldog, Ference
; APPLICANT: Smithson, Glenda
; APPLICANT: Rastelli, Luca
TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYPEPTIDES ENCODING THEM AND
FILE REFERENCE: CURA-65 CIP
CURRENT FILING DATE: 2002-11-26
PRIOR APPLICATION NUMBER: US/10/098,871
PRIOR FILING DATE: 2000-09-12
PRIOR APPLICATION NUMBER: 60/153,629
PRIOR FILING DATE: 1999-09-13
PRIOR APPLICATION NUMBER: 60/154,520
PRIOR FILING DATE: 1999-09-16
PRIOR APPLICATION NUMBER: 60/154,762
PRIOR FILING DATE: 1999-09-20
PRIOR APPLICATION NUMBER: 60/159,231
PRIOR FILING DATE: 2000-10-31
PRIOR APPLICATION NUMBER: 60/276,960
PRIOR FILING DATE: 2001-03-19
NUMBER OF SEQ ID NOS: 80
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 65
LENGTH: 1056
TYPE: PRT
ORGANISM: Xenopus laevis
US-10-098-871-65

Query Match 100.0%; Score 32; DB 12; Length 1056;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KXKQPG 6
DB 625 KXKQPG 630

RESULT 2

US-09-764-864-1162
; Sequence 1162, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1162
; LENGTH: 319
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (3)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (23)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (196)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; US-09-764-864-1162

Query Match 93.8%; Score 30; DB 10; Length 319;
Best Local Similarity 83.3%; Pred. No. 2.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KXKQPG 6
DB 93 KXKQPG 98

RESULT 3

US-10-054-935-2
; Sequence 2, Application US/10054935
; Publication No. US20030143546A1
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies, Inc
; TITLE OF INVENTION: BREAST CANCER TRANSCRIPTION FACTOR GENE AND USES
; FILE REFERENCE: 16U 107 R1
; CURRENT APPLICATION NUMBER: US/10/054,935
; CURRENT FILING DATE: 2002-01-25
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 614
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-054-935-2

Query Match 90.6%; Score 29; DB 12; Length 614;
Best Local Similarity 83.3%; Pred. No. 7.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KXKQPG 6
DB 51 KXKQPG 56

RESULT 4

US-10-144-194A-66
; Sequence 66, Application US/10144194A
; Publication No. US20030215809A1
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies Inc
; TITLE OF INVENTION: Regulated Breast Cancer Genes
; FILE REFERENCE: 3U 103 R1
; CURRENT APPLICATION NUMBER: US/10/144,194A
; CURRENT FILING DATE: 2002-06-12
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 66
; LENGTH: 614
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-144-194A-66

Query Match 90.6%; Score 29; DB 12; Length 614;
Best Local Similarity 83.3%; Pred. No. 7.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KXKQPG 6
DB 51 KXKQPG 56

RESULT 5

US-10-054-935-4
; Sequence 4, Application US/10054935
; Publication No. US20030143546A1
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies, Inc
; TITLE OF INVENTION: BREAST CANCER TRANSCRIPTION FACTOR GENE AND USES
; FILE REFERENCE: 16U 107 R1
; CURRENT APPLICATION NUMBER: US/10/054,935
; CURRENT FILING DATE: 2002-01-25
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 616
; TYPE: PRT
; ORGANISM: Mus musculus
; US-10-054-935-4

Query Match 90.6%; Score 29; DB 12; Length 616;
Best Local Similarity 83.3%; Pred. No. 7.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KXKQPG 6
DB 51 KXKQPG 56

RESULT 6

US-10-354-358-4
; Sequence 4, Application US/10354358
; Publication No. US20030157082A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc
; APPLICANT: Hunter, John Joseph
; APPLICANT: Macbeth, Kyle J.
; APPLICANT: Tsai, Fong-Ying
; APPLICANT: Leeson, Andrea
; APPLICANT: Lightcap, Eric S.
; APPLICANT: Williamson, Mark
; APPLICANT: Rudolph-Owen, Laura A.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
; TITLE OF INVENTION: CANCER USING 140, 1470, 1686, 2089, 2427, 3702, 5891, 6428,
; TITLE OF INVENTION: 7181, 7660, 25641, 69583, 49863, 8897, 1682, 17677, 9235,
; TITLE OF INVENTION: 3703, 14171, 10359, 1660, 1450, 18894, 2088, 32427, 2160,
; TITLE OF INVENTION: 9252, 9389, 1642, 85269, 10297, 1584, 9525, 14124, 4469,

TITLE OF INVENTION: 8990, 2100, 9288, 64698, 10480, 20893, 33230, 1586, 9943,
TITLE OF INVENTION: 16334, 68862, 9011, 14031, 6178, 21225, 1420, 32236, 2099,
FILE REFERENCE: MP102-0201R00N01M
CURRENT APPLICATION NUMBER: US/10/354,358
PRIOR FILING DATE: 2003-01-30
PRIOR APPLICATION NUMBER: US 60/353,600
PRIOR FILING DATE: 2002-01-31
PRIOR APPLICATION NUMBER: US 60/364,517
PRIOR FILING DATE: 2002-03-15
PRIOR APPLICATION NUMBER: US 60/371,075
PRIOR FILING DATE: 2002-04-09
PRIOR APPLICATION NUMBER: US 60/371,507
PRIOR FILING DATE: 2002-04-10
PRIOR APPLICATION NUMBER: US 60/372,984
PRIOR FILING DATE: 2002-04-16
PRIOR APPLICATION NUMBER: US 60/374,194
PRIOR FILING DATE: 2002-04-19
PRIOR APPLICATION NUMBER: US 60/382,995
PRIOR FILING DATE: 2002-05-24
PRIOR APPLICATION NUMBER: US 60/385,023
PRIOR FILING DATE: 2002-05-31
PRIOR APPLICATION NUMBER: US 60/388,853
PRIOR FILING DATE: 2002-06-14
PRIOR APPLICATION NUMBER: US 60/389,395
PRIOR FILING DATE: 2002-06-17
Remaining Prior Application data removed - See File Wrapper or PALM.
SOFTWARE: Seq ID NOS: 122
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 4
LENGTH: 998
TYPE: PRT
ORGANISM: Homo sapiens
US-10-354-358-4

Query Match 90.6%; Score 29; DB 12; Length 998;
Best Local Similarity 83.3%; Pred. No. 1.2e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKKPG 6
Db 651 RKKPG 656

RESULT 7
US-09-056-160B-9
Sequence 9, Application US/09056160B
Patent No. US20020032315A1
GENERAL INFORMATION:
APPLICANT: Baca, Manuel
APPLICANT: Wells, James A.
APPLICANT: Presta, Leonard G.
APPLICANT: Lowman, Henry B.
APPLICANT: Chen, Yvonne M.
TITLE OF INVENTION: ANTI-VEGF ANTIBODIES
NUMBER OF SEQUENCES: 131
CORRESPONDENCE ADDRESS:
ADDRESSER: Genentech, Inc.
STREET: 1 DNA Way
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Winpatin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/056,160B
FILING DATE: 06-Apr-1998
CLASSIFICATION: 424
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/054,856
FILING DATE: 06-AUG-1997
ATTORNEY/AGENT INFORMATION:
NAME: Hasak, Janet E.
REGISTRATION NUMBER: 28,616
REFERENCE/DOCKET NUMBER: P1093R2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-1896
TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 123 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
US-09-056-160B-9

Query Match 87.5%; Score 28; DB 9; Length 123;
Best Local Similarity 83.3%; Pred. No. 2.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKKPG 6
Db 10 RKKPG 15

RESULT 8
US-10-234-671-9
Sequence 9, Application US/10234671
Publication No. US20030190317A1
GENERAL INFORMATION:
APPLICANT: Baca, Manuel
APPLICANT: Wells, James A.
APPLICANT: Presta, Leonard G.
APPLICANT: Lowman, Henry B.
APPLICANT: Chen, Yvonne M.
TITLE OF INVENTION: ANTI-VEGF ANTIBODIES
NUMBER OF SEQUENCES: 131
CORRESPONDENCE ADDRESS:
ADDRESSER: Genentech, Inc.
STREET: 1 DNA Way
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Winpatin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/234,671
FILING DATE: 03-Sep-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/056160
FILING DATE: 06-Apr-1998
APPLICATION NUMBER: 60/126446
FILING DATE: 07-Apr-1997
APPLICATION NUMBER: 60/054856
FILING DATE: 06-AUG-1997
ATTORNEY/AGENT INFORMATION:
NAME: Cui, Steven X.
REGISTRATION NUMBER: 44,637
REFERENCE/DOCKET NUMBER: P1093R2C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-8674
TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 123 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
SEQUENCE DESCRIPTION: SEQ ID NO: 9:

US-10-234-671-9

Query Match 87.5%; Score 28; DB 12; Length 123;
Best Local Similarity 83.3%; Pred. No. 2.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLRQPG 6
:|||||
Db 10 ELKQPG 15

RESULT 9

US-09-738-626-5322
; Sequence 5322, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: MAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHITO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738, 626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 5322
; LENGTH: 185
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-5322

Query Match 87.5%; Score 28; DB 10; Length 185;
Best Local Similarity 83.3%; Pred. No. 3.6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLRQPG 6
:|||||
Db 58 KLRQPG 63

RESULT 10
US-09-880-748-1891
; Sequence 1891, Application US/09880748
; Publication No. US2003005937A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
; FILE REFERENCE: PF523
; CURRENT APPLICATION NUMBER: US/09/880, 748
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/212, 210
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 60/240, 816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/276, 248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/277, 379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/293, 499
; PRIOR FILING DATE: 2001-05-25

; NUMBER OF SEQ ID NOS: 3239
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1891
; LENGTH: 249
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-748-1891

Query Match 87.5%; Score 28; DB 11; Length 249;
Best Local Similarity 83.3%; Pred. No. 4.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLRQPG 6
:|||||
Db 10 ELKQPG 15

RESULT 11
US-09-880-748-1894
; Sequence 1894, Application US/09880748
; Publication No. US2003005937A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
; FILE REFERENCE: PF523
; CURRENT APPLICATION NUMBER: US/09/880, 748
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/212, 210
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 60/240, 816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/276, 248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/277, 379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/293, 499
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 3239
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1894
; LENGTH: 249
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-748-1894

Query Match 87.5%; Score 28; DB 11; Length 249;
Best Local Similarity 83.3%; Pred. No. 4.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLRQPG 6
:|||||
Db 10 ELKQPG 15

RESULT 12
US-09-880-748-1904
; Sequence 1904, Application US/09880748
; Publication No. US2003005937A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
; FILE REFERENCE: PF523
; CURRENT APPLICATION NUMBER: US/09/880, 748
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/212, 210
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 60/240, 816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/276, 248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/277, 379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/293, 499

PRIOR FILING DATE: 2001-05-25
NUMBER OF SEQ ID NOS: 3239
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1904
LENGTH: 249
TYPE: PRT
ORGANISM: Homo sapiens
US-09-880-748-1904

Query Match 87.5%; Score 28; DB 11; Length 249;
Best Local Similarity 83.3%; Pred. No. 4.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KXKOPG 6
:|||||
Db 10 ELKOPG 15

RESULT 13
US-09-880-748-2085
Sequence 2085, Application US/09880748
Publication No. US2003005937A1
GENERAL INFORMATION:
APPLICANT: Ruben et al.
TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
FILE REFERENCE: PFS23
CURRENT APPLICATION NUMBER: US/09/880,748
CURRENT FILING DATE: 2001-06-15
PRIOR APPLICATION NUMBER: 60/212,210
PRIOR FILING DATE: 2000-06-15
PRIOR APPLICATION NUMBER: 60/240,816
PRIOR FILING DATE: 2000-10-17
PRIOR APPLICATION NUMBER: 60/276,248
PRIOR FILING DATE: 2001-03-16
PRIOR APPLICATION NUMBER: 60/277,379
PRIOR FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/293,499
PRIOR FILING DATE: 2001-05-25
NUMBER OF SEQ ID NOS: 3239
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2085
LENGTH: 249
TYPE: PRT
ORGANISM: Homo sapiens
US-09-880-748-2085

Query Match 87.5%; Score 28; DB 11; Length 249;
Best Local Similarity 83.3%; Pred. No. 4.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KXKOPG 6
:|||||
Db 10 ELKOPG 15

RESULT 14
US-09-880-748-2120
Sequence 2120, Application US/09880748
Publication No. US2003005937A1
GENERAL INFORMATION:
APPLICANT: Ruben et al.
TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
FILE REFERENCE: PFS23
CURRENT APPLICATION NUMBER: US/09/880,748
CURRENT FILING DATE: 2001-06-15
PRIOR APPLICATION NUMBER: 60/212,210
PRIOR FILING DATE: 2000-06-15
PRIOR APPLICATION NUMBER: 60/240,816
PRIOR FILING DATE: 2000-10-17
PRIOR APPLICATION NUMBER: 60/276,248
PRIOR FILING DATE: 2001-03-16
PRIOR APPLICATION NUMBER: 60/277,379
PRIOR FILING DATE: 2001-03-21

PRIOR APPLICATION NUMBER: 60/293,499
PRIOR FILING DATE: 2001-05-25
NUMBER OF SEQ ID NOS: 3239
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2120
LENGTH: 249
TYPE: PRT
ORGANISM: Homo sapiens
US-09-880-748-2120

Query Match 87.5%; Score 28; DB 11; Length 249;
Best Local Similarity 83.3%; Pred. No. 4.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KXKOPG 6
:|||||
Db 10 ELKOPG 15

RESULT 15
US-09-880-748-2122
Sequence 2122, Application US/09880748
Publication No. US2003005937A1
GENERAL INFORMATION:
APPLICANT: Ruben et al.
TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
FILE REFERENCE: PFS23
CURRENT APPLICATION NUMBER: US/09/880,748
CURRENT FILING DATE: 2001-06-15
PRIOR APPLICATION NUMBER: 60/212,210
PRIOR FILING DATE: 2000-06-15
PRIOR APPLICATION NUMBER: 60/240,816
PRIOR FILING DATE: 2000-10-17
PRIOR APPLICATION NUMBER: 60/276,248
PRIOR FILING DATE: 2001-03-16
PRIOR APPLICATION NUMBER: 60/277,379
PRIOR FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/293,499
PRIOR FILING DATE: 2001-05-25
NUMBER OF SEQ ID NOS: 3239
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2122
LENGTH: 249
TYPE: PRT
ORGANISM: Homo sapiens
US-09-880-748-2122

Query Match 87.5%; Score 28; DB 11; Length 249;
Best Local Similarity 83.3%; Pred. No. 4.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KXKOPG 6
:|||||
Db 10 ELKOPG 15

Search completed: December 23, 2003, 15:38:40
Job time : 12.25 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 23, 2003, 15:33:08 ; Search time 6.75 Seconds
(without alignments)
85.483 Million cell updates/sec

Title: US-09-980-564-12

Perfect score: 32

Sequence: 1 KXKOPG 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	32	100.0	171	2	A29795
2	32	100.0	281	2	AB3313
3	32	100.0	343	2	A29319
4	32	100.0	367	2	A32068
5	32	100.0	386	2	A48571
6	32	100.0	387	2	D41156
7	32	100.0	387	2	C41156
8	32	100.0	388	2	A39756
9	32	100.0	412	1	OZKOAF
10	32	100.0	424	2	A54533
11	32	100.0	442	2	A54529
12	30	93.8	277	2	G83928
13	29	90.6	260	1	JS0341
14	29	90.6	280	2	AF3235
15	29	90.6	470	2	T32107
16	29	90.6	737	2	H55882
17	29	90.6	998	2	S37627
18	28	87.5	299	2	E69288
19	28	87.5	299	2	H69462
20	28	87.5	357	2	A69426
21	28	87.5	379	1	YWBVM
22	28	87.5	466	2	AD2537
23	28	87.5	475	2	C87567
24	28	87.5	502	2	D86765
25	28	87.5	508	2	A62525
26	28	87.5	550	2	A46419
27	28	87.5	630	2	A70117
28	28	87.5	686	2	A70104
29	28	87.5	695	2	I54325

ALIGNMENTS

30	28	87.5	1938	1	A40997	myosin heavy chain
31	28	87.5	2022	2	T48818	glucan 1,4-alpha-g
32	28	87.5	3712	2	S18253	lamnin alpha-1 ch
33	27	84.4	62	2	E82790	hypothetical prote
34	27	84.4	107	2	S69077	hypothetical prote
35	27	84.4	115	2	H75086	hypothetical prote
36	27	84.4	121	2	D86468	hypothetical prote
37	27	84.4	143	2	AE2221	hypothetical prote
38	27	84.4	164	2	G45045	hypocetylthrin I
39	27	84.4	164	2	S43779	phycocerythrin I
40	27	84.4	198	2	S53032	hypocetylthrin I
41	27	84.4	211	2	AC1573	Salmonella enteric
42	27	84.4	211	2	AH1219	Salmonella enteric
43	27	84.4	213	2	C66575	hypocetylthrin I
44	27	84.4	213	2	B72045	hypocetylthrin I
45	27	84.4	218	2	AH2758	conserved hypothet

RESULT 1

A29795
circumsporozoite protein - malaria parasite (Plasmodium falciparum) (strain B1) (Eragme
C:Species: Plasmodium falciparum
C:Date: 28-Aug-1989 #sequence_revision 28-Aug-1989 #text_change 09-Jun-2000
C:Accession: A29795

Ride la Cruz, V.F.; Lal, A.A.; McCutchan, T.F.

U. Biol. Chem. 262, 11935-11939, 1987

A:Title: Sequence variation in putative functional domains of the circumsporozoite prote

A:Reference number: A92609; MUID:87308186; PMID:2442154

A:Accession: A29795

A:Molecule type: DNA

A:Residues: 1-171 <DELA>

C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology

Query Match

Best Local Similarity 100.0%; Score 32; DB 2; Length 171;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KXKOPG 6
DB 105 KXKOPG 110

RESULT 2

AB3313
proliporotein diacylglycerol transferase (EC 2.4.99.-) [imported] - Brucella melitensis

C:Species: Brucella melitensis

C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 15-Feb-2002

C:Accession: AB3313

R:DelVecchio, V.G.; Kapatal, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova, '.

Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002

A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitens

A:Reference number: AD3252; PMID:11756688

A:Accession: AB3313

A:Molecule type: DNA

A:Residues: 1-281 <KUR>

A:Cross-references: GB:AB008917; PIDN:AMLS1669.1; PID:G17982400; GSPDB:GN00190

A:Experimental source: strain 16M

C:Genetics:

A:Gene: BMEI0488

A:Map position: 1

C:Superfamily: proliporotein diacylglycerol transferase

C:Keywords: glycosyltransferase

Query Match 100.0%; Score 32; DB 2; Length 281;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KXKOPG 6

Db 212 KLKOPG 217

RESULT 3

A29319

circumsporozoite protein - Plasmodium vivax (strain Sal-1) (fragment)

C/Species: Plasmodium vivax

C/Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 17-Mar-1999

C/Accession: A29319; A44187

R/De la Cruz, V.F.; Lal, A.A.; Welsh, J.A.; McCutchan, T.F.

J. Biol. Chem. 262, 6464-6467, 1987

A/Title: Evolution of the immunodominant domain of the circumsporozoite protein gene from

A/Reference number: A29319; PMID:87194878; PMID:2437120

A/Accession: A29319

A/Molecule type: DNA

A/Residues: 1-343

R/McCutchan, T.F.; Lal, A.A.; de la Cruz, V.F.; Miller, L.H.; Maloy, W.L.; Charoenvit, Y

Science 230, 1381-1383, 1985

A/Title: Sequence of the immunodominant epitope for the surface protein on sporozoites of

A/Reference number: A44187; PMID:86070222; PMID:2416057

A/Accession: A44187

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-98 <MCC>

A/Cross-references: GB:J02751

C/Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology

F;268-321/Domain: thrombospondin type 1 repeat homology <THRI>

Query Match 100.0%; Score 32; DB 2; Length 343;

Best Local Similarity 100.0%; Pred. No. 18;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLKOPG 6

Db 57 KLKOPG 62

RESULT 4

A32068

circumsporozoite protein - Plasmodium vivax (strain North Korean) (fragment)

C/Species: Plasmodium vivax

C/Date: 21-May-1990 #sequence_revision 21-May-1990 #text_change 20-Aug-1999

C/Accession: A32068

R/Arnott, D.E.; Barnwell, J.W.; Stewart, M.J.

Proc. Natl. Acad. Sci. U.S.A. 85, 8102-8106, 1988

A/Title: Does biased gene conversion influence polymorphism in the circumsporozoite prot

A/Reference number: A32068; PMID:89042133; PMID:3054880

A/Accession: A32068

A/Molecule type: DNA

A/Residues: 1-367 <ARN>

A/Cross-references: GB:M20670; GB:J04090; NID:G160183; PIDN:AAA29534.1; PID:9552194

C/Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology

F;1292-345/Domain: thrombospondin type 1 repeat homology <THRI>

Query Match 100.0%; Score 32; DB 2; Length 367;

Best Local Similarity 100.0%; Pred. No. 20;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLKOPG 6

Db 65 KLKOPG 70

RESULT 5

A48571

circumsporozoite protein PVCS type 1 - Plasmodium simium

C/Species: Plasmodium simium

C/Date: 01-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 20-Aug-1999

C/Accession: A48571

R/Goldman, I.F.; Qari, S.H.; Miller, P.G.; Collins, W.E.; Lal, A.A.

Mol. Biochem. Parasitol. 57, 177-180, 1993

A/Title: Circumsporozoite protein gene of Plasmodium simium, a Plasmodium vivax-like mon

RESULT 8

A39756
circumsporozoite protein - Plasmodium reichenowi
C/Species: Plasmodium reichenowi
C/Date: 14-Feb-1992 #sequence_revision 14-Feb-1992 #text_change 20-Aug-1999
C/Accession: A39756
R/Ref: A.A.; Goldman, I.F.
J. Biol. Chem. 266, 6686-6689, 1991
A/Title: Circumsporozoite protein gene from Plasmodium reichenowi, a chimpanzee malaria
A/Reference number: A39756; MUID:91201303; PMID:2016283
A/Accession: A39756
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-388 <DAL>
A/Cross-references: GB:M60972; NID:g160228; PIDN:AAA29561.1; PID:g160229
C/Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
F/312-366/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match 100.0%; Score 32; DB 2; Length 388;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KIKOPG 6
Db 112 KIKOPG 117

RESULT 9

OZZOAF
circumsporozoite protein - malaria parasite (Plasmodium falciparum) (isolate INTM22)
C/Species: Plasmodium falciparum
C/Date: 15-Nov-1984 #sequence_revision 15-Nov-1984 #text_change 09-Jun-2000
C/Accession: A03388
R/Ref: J.B.; Williams, J.L.; McCutchan, T.F.; Weber, J.L.; Wirtz, R.A.; Hochmeyer, W.T.
Science 225, 593-599, 1984
A/Title: Structure of the gene encoding the immunodominant surface antigen on the sporozoite
A/Reference number: A03388; MUID:84250215; PMID:6204383
A/Accession: A03388
A/Molecule type: DNA
A/Residues: 1-412 <DAM>
A/Cross-references: GB:K02194; NID:g160160; PIDN:AAA29524.1; PID:g160161
A/Experimental source: clone 7G8
C/Comment: Residues 1-16 are the probable signal sequence.
C/Comment: There are 41 copies of a 4-residue repeating unit in the middle domain of the
F/336-390/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match 100.0%; Score 32; DB 1; Length 412;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KIKOPG 6
Db 112 KIKOPG 117

RESULT 10

A54533
circumsporozoite protein - malaria parasite (Plasmodium falciparum) (strain T4, Thailand)
C/Species: Plasmodium falciparum
C/Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 09-Jun-2000
C/Accession: A54533
R/Ref: Portillo, H.A.; Nussenzweig, R.S.; Dnea, V.
Mol. Biochem. Parasitol. 24, 289-294, 1987
A/Title: Circumsporozoite protein of a Plasmodium falciparum strain from Thailand.
A/Reference number: A54533; MUID:87315205; PMID:3306373
A/Accession: A54533
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-424
A/Cross-references: GB:M19752; NID:g160216; PIDN:AAA29555.1; PID:g160217
C/Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
F/346-402/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match 93.8%; Score 30; DB 2; Length 277;
Best Local Similarity 83.3%; Pred. No. 43;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Query Match 100.0%; Score 32; DB 2; Length 424;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KIKOPG 6
Db 112 KIKOPG 117

RESULT 11

A54529
circumsporozoite protein - malaria parasite (Plasmodium falciparum) (strain wellcome)
C/Species: Plasmodium falciparum
C/Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 09-Jun-2000
C/Accession: A54529
R/Ref: Lockyer, M.J.; Schwarz, R.T.
Mol. Biochem. Parasitol. 22, 101-108, 1987
A/Title: Strain variation in the circumsporozoite protein gene of Plasmodium falciparum
A/Reference number: A54529; MUID:87115616; PMID:3543671
A/Accession: A54529
A/Status: preliminary; not compared with conceptual translation
A/Molecule type: DNA
A/Residues: 1-442 <LOC>
A/Cross-references: GB:M15505; NID:g160214; PIDN:AAA29554.1; PID:g160215
C/Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
C/Keywords: tandem repeat
F/366-420/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match 100.0%; Score 32; DB 2; Length 442;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KIKOPG 6
Db 122 KIKOPG 127

RESULT 12

GB3908
hypothetical protein BH2071 [imported] - Bacillus halodurans (strain C-125)
C/Species: Bacillus halodurans
C/Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C/Accession: GB3908
R/Ref: Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A/Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A/Reference number: A83650; MUID:20512582; PMID:11058132
A/Accession: GB3908
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-277 <STO>
A/Cross-references: GB:AP001514; GB:BA000004; NID:g10174613; PIDN:BA05790.1; GSPDB:GN00
A/Experimental source: strain C-125
C/Genetics:
A/Gene: BH2071

Query Match 93.8%; Score 30; DB 2; Length 277;
Best Local Similarity 83.3%; Pred. No. 43;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KIKOPG 6
Db 237 KIKOPG 242

RESULT 13

JS0341
indole-3-glycerol-phosphate synthase (EC 4.1.1.48) - Lactobacillus casei
N/Alternate names: trpC protein
C/Species: Lactobacillus casei
C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000
C/Accession: S42344; JS0341

R;Natori, Y.; Kano, Y.; Imamoto, F.
 J. Biochem. 107, 248-255, 1990
 A;Title: Nucleotide sequences and genomic constitution of five tryptophan genes of *Lactococcus lactis*
 A;Reference number: S42342; MUID:90299861; PMID:2113923
 A;Accession: S42344
 A;Molecule type: DNA
 A;Residues: 1-260 <NNT>
 A;Cross-references: EMBL:D00496; NID:g216754; PIDD:BA00384.1; PID:g216757
 A;Experimental source: isolate RNL7
 C;Genetics:
 A;Gene: trpC
 C;Superfamily: indole-3-glycerol-phosphate synthase; trpC homology
 C;Keywords: carbon-carbon lyase; carboxy-lyase; tryptophan biosynthesis
 F;2-256/Domain: trpC homology <TRC>

Query Match 90.6%; Score 29; DB 1; Length 260;
 Best Local Similarity 83.3%; Pred. No. 68;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLRPG 6
 |||||
 DB 44 KLRPG 49

RESULT 14

AF3235

haloalkane dehalogenase [imported] - *Agrobacterium tumefaciens* (strain C58, Dupont) plasmidC;Species: *Agrobacterium tumefaciens*

C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002

C;Accession: AF3235

R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.; Karp, P.; Romero, P.; Zhang, S.

Science 294, 2317-2323, 2001

A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Kreppan, W.; Perry, M.; Gordon-Kamm, S.; E.W.

A;Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.

A;Reference number: AB2577; MUID:21608550; PMID:11743193

A;Accession: AF3235

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-280 <KUR>

A;Cross-references: GB:AE008690; PIDD:AAL46300.1; PID:g17744084; GSPDB:GN00189

A;Experimental source: strain C58 (Dupont)

C;Genetics:

A;Gene: dha

A;Genome: plasmid

Query Match 90.6%; Score 29; DB 2; Length 280;
 Best Local Similarity 83.3%; Pred. No. 73;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLRPG 6
 |||||

DB 135 KLRPG 140

RESULT 15

T32107

hypothetical protein F45C12.3 - *Caenorhabditis elegans*C;Species: *Caenorhabditis elegans*

C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999

C;Accession: T32107

R;Johnson, D.; Kramer, J.; Keppeler, D.

submitted to the EMBL Data Library, July 1997

A;Description: The sequence of *C. elegans* cosmid F45C12.

A;Reference number: Z21123

A;Accession: T32107

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-470 <TOH>

A;Cross-references: EMBL:AF016884; PIDD:AA866203.1; GSPDB:GN00020; CBSP:F45C12.3

A;Experimental source: strain Bristol N2; clone F45C12

C;Genetics:
 A;Gene: CBSP:F45C12.3
 A;Map position: 2
 A;Introns: 20/3; 80/3; 162/3

Query Match 90.6%; Score 29; DB 2; Length 470;
 Best Local Similarity 83.3%; Pred. No. 126+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLRPG 6
 |||||
 DB 306 KLRPG 311

Search completed: December 23, 2003, 15:34:24
 Job time : 7.75 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 23, 2003, 15:33:08 ; Search time 5 Seconds
(without alignments)
56.432 Million cell updates/sec

Title: US-09-980-564-12

Perfect score: 32

Sequence: 1 KLKOPG 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Swissprot_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	32	100.0	343	1 CSP_PLAVS	P13826 plasmodium
2	32	100.0	378	1 CSP_PLAVB	P08677 plasmodium
3	32	100.0	386	1 CSP_PLARI	Q03110 plasmodium
4	32	100.0	388	1 CSP_PLARI	P26694 plasmodium
5	32	100.0	412	1 CSP_PLARA	P03893 plasmodium
6	32	100.0	424	1 CSP_PLAPF	P13814 plasmodium
7	32	100.0	442	1 CSP_PLAPW	P08307 plasmodium
8	30	93.8	1186	1 PKCB_HUMAN	Q94104 homo sapien
9	29	90.6	260	1 TRPC_LACCA	P17217 lactobacilli
10	29	90.6	304	1 DHAA_AGRIS	Q84671 agrobacteri
11	29	90.6	304	1 EPH3_HUMAN	P54753 homo sapien
12	29	90.6	1083	1 T2D3_HUMAN	P00268 homo sapien
13	28	87.5	379	1 SYMW_YEAST	P04803 saccharomyc
14	28	87.5	588	1 HDAP_MOUSE	Q09413 mus musculu
15	28	87.5	695	1 XET_HUMAN	Q02040 homo sapien
16	28	87.5	1011	1 HDAP_HUMAN	Q94104 homo sapien
17	28	87.5	1938	1 MYS_ABOIR	P24733 aequipecten
18	28	87.5	3712	1 LMA_DEOME	Q00174 dirosophila
19	27	84.4	107	1 YP94_YEAST	Q06835 saccharomyc
20	27	84.4	164	1 PHAI_SYNPW	Q08086 synchococc
21	27	84.4	164	1 PHAI_SYNPW	Q02179 synchococc
22	27	84.4	198	1 YMO3_YEAST	Q03673 saccharomyc
23	27	84.4	357	1 CAD2_ARATH	Q04882 arabidopsi
24	27	84.4	417	1 ODO2_BACSU	P16263 bacillus su
25	27	84.4	425	1 YTRI_EBV	P30119 Epstein-Bar
26	27	84.4	452	1 ODO2_SCHPO	Q04681 schizosach
27	27	84.4	456	1 ENO_MYCPU	Q09450 mycoplasma
28	27	84.4	486	1 GAG_EIAVY	P03351 equine infe
29	27	84.4	588	1 FRD2_SHEER	Q94400 shewanella
30	27	84.4	594	1 UVRC_HELPJ	Q94121 helicobacte
31	27	84.4	594	1 UVRC_HELPJ	P56428 helicobacte
32	27	84.4	701	1 PALI_ORYSA	P14717 oryza sativ
33	27	84.4	708	1 ATB2_ANASP	Q945ds anabaena sp

34	27	84.4	833	1 MAK1_HUMAN	Q92918 homo sapien
35	27	84.4	844	1 AMPN_LACBE	Q10730 lactobacilli
36	27	84.4	854	1 UN33_CAEEL	Q01630 caenorhabdi
37	27	84.4	1812	1 BRCL_MOUSE	P48754 mus musculu
38	27	84.4	2233	1 COAC_YEAST	Q00955 saccharomyc
39	26	81.2	65	1 RL29_BUCAT	P57583 buchnera ap
40	26	81.2	65	1 RL29_BUCAP	Q8K958 buchnera ap
41	26	81.2	97	1 PLGM_ECOLI	P43532 escherichia
42	26	81.2	97	1 PLGM_SALTY	P26477 salmonella
43	26	81.2	99	1 PLGM_YEREN	Q57401 yersinia en
44	26	81.2	152	1 NDK_NEUCR	Q94uy8 neurospora
45	26	81.2	153	1 YM38_MYCTU	Q10520 mycobacteri

ALIGNMENTS

RESULT 1
ID CSP_PLAVS STANDARD; PRT; 343 AA.
AC P13826;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Circumsporozoite protein (CS) (Fragment).
OS Plasmodium vivax (strain Salvador I).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=126793;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86070222; PubMed=2416057;
RA McCutchan T.F., Lal A.A., de la Cruz V.F., Miller L.H., Maloy W.L.,
RA Chroenavit Y., Beaudoin R.L., Guerry P., Wistar R. Jr., Hoffman S.L.,
RA Hockmeyer W.T., Collins W.E., Wirth D.,
RT "Sequence of the immunodominant epitope for the surface protein on
RT sporozoites of Plasmodium vivax."
RL Science 230:1381-1383 (1985).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=87194878; PubMed=2437120;
RA de la Cruz V.F., Lal A.A., Welsh J.A., McCutchan T.F.,
RT "Evolution of the immunodominant domain of the circumsporozoite
RT protein gene from Plasmodium vivax. Implications for vaccines."
RL J. Biol. Chem. 262:6464-6467 (1987).
CC - FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
CC SURFACE ANTIGEN ON THE SPOOROZOITE (THE INFECTIVE STAGE OF THE
CC MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
CC VERTEBRATE HOST).
CC - MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
CC ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
CC WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC - SIMILARITY: Contains 1 TSP type-1 domain.
CC InterPro; IPR003067; Circmprozoite.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; CIRCSPOROZITE.
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PS50092; TSP1; 1.
FT Sporozoite; Malaria; Repeat.
FT NON TER
FT DOMAIN 1
FT 63 243
FT 19 X 9 AA TANDM REPEATS OF P-G-D-R-A-D-
FT G-O-P
FT TSP-TYPE-1
SQ SEQUENCE 343 AA; 3415 MW; 308EFD5B3C15DFC3 CRC64;
Query Match 100.0%; Score 32; DB 1; Length 343;
Best Local Similarity 100.0%; Pred. No. 5.3;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

RESULT 2
CSP_PLAVB
ID - CSP_PLAVB      STANDARD;      PRT;      378 AA.
AC P08677;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Circumsporozoite protein precursor (CS).
OS Plasmodium vivax (strain Belem).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
NCBI_TaxID=31273;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86044510; PubMed=2414847;
RA Arnot D.E., Barnwell J.W., Tam J.P., Nussenzweig V., Nussenzweig R.S.,
RT "Circumsporozoite protein of Plasmodium vivax: gene cloning and
RT characterization of the immunodominant epitope."
RL Science 230:815-818 (1985).
RN [2]
RP REVISIONS.
RX MEDLINE=89042133; PubMed=3054880;
RA Arnot D.E., Barnwell J.W., Stewart M.J.;
RT "Does biased gene conversion influence polymorphism in the
RT circumsporozoite protein-encoding gene of Plasmodium vivax?";
RL Proc. Natl. Acad. Sci. U.S.A. 85:8102-8106 (1988).
CC -1- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
CC SURFACE ANTIGEN ON THE SPOROZOITE (THE INFECTIVE STAGE OF THE
CC MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
CC VERTEBRATE HOST).
CC -1- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
CC ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
CC WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -1- SIMILARITY: Contains 1 TSP type-1 domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M1926; AAA29526.1; -.
DR EMBL; J02751; AAA29529.1; ALT_SEQ.
DR PIR; A26256; OZZOAV.
DR InterPro; IPR003067; Circmsporzoite.
DR InterPro; IPR000884; TSPL.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSPL; 1.
DR PROSITE; PS50092; TSPL; 1.
DR Malaria; Sporozoite; Repeat; Signal.
FT SIGNAL 1 19
FT CHAIN 20 378
FT DOMAIN 97 267
FT -----
FT DOMAIN 304 356
FT CONFLICT 36 36 G->E (IN REF. 1).
FT CONFLICT 96 96 G->R (IN REF. 1).
FT CONFLICT 295 295 E->A (IN REF. 1).
FT CONFLICT 328 328 R->S (IN REF. 1).
SQ SEQUENCE 378 AA; 37800 MW; C84B5BED05B3C9ED CRC64;

Query Match 100.0%; Score 32; DB 1; Length 378;
Best Local Similarity 100.0%; Pred. No. 6;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KTKOPG 6
DB 91 KTKOPG 96

```

```

RESULT 3
CSP_PLASI
ID - CSP_PLASI      STANDARD;      PRT;      386 AA.
AC Q0310;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Circumsporozoite protein precursor (CS).
GN CS.
OS Plasmodium simium.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
NCBI_TaxID=5859;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93149205; PubMed=8426613;
RA Goldman I.F., Qari S.H., Millet P.G., Collins W.E., Lal A.A.;
RT "Circumsporozoite protein gene of Plasmodium simium, a Plasmodium
RT vivax-like monkey malaria parasite."
RL Mol. Biochem. Parasitol. 57:177-180 (1993).
CC -1- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
CC SURFACE ANTIGEN ON THE SPOROZOITE (THE INFECTIVE STAGE OF THE
CC MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
CC VERTEBRATE HOST).
CC -1- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
CC ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
CC WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -1- SIMILARITY: Contains 1 TSP type-1 domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; I05068; AAA29525.1; -.
DR PIR; A48571; A48571.
DR InterPro; IPR003067; Circmsporzoite.
DR InterPro; IPR000884; TSPL.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSPL; 1.
DR PROSITE; PS50092; TSPL; 1.
DR Malaria; Sporozoite; Signal; Repeat.
FT SIGNAL 1 19
FT CHAIN 20 386
FT DOMAIN 96 275
FT -----
FT DOMAIN 312 364
FT CONFLICT 364 364 TSP TYPE-1.
SQ SEQUENCE 386 AA; 38567 MW; A0097DA4DE5548DB CRC64;

Query Match 100.0%; Score 32; DB 1; Length 386;
Best Local Similarity 100.0%; Pred. No. 6.1;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KTKOPG 6
DB 91 KTKOPG 96

```

```

RESULT 4
CSP_PLARE
ID - CSP_PLARE      STANDARD;      PRT;      388 AA.
AC P26694;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Circumsporozoite protein precursor (CS).
OS Plasmodium reichenowi.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
NCBI_TaxID=5854;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91201303; PubMed=2016283;
RA Lal A.A., Goldman I.F.;
RT "Circumsporozoite protein gene from Plasmodium reichenowi, a
RT malaria parasite Plasmodium falciparum."
RL J. Biol. Chem. 266:6686-6689(1991).
CC -1- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
CC SURFACE ANTIGEN ON THE SPOOROZOITE (THE INFECTIVE STAGE OF THE
CC MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
CC VERTEBRATE HOST).
CC -1- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
CC ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
CC WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -1- SIMILARITY: Contains 1 TSP type-1 domain.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; M60972; AAA29561.1; -.
DR PIR; A39756; A39756.
DR InterPro; IPR003067; Circmsprzoite.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PRO1303; CRCMSPRZOITE.
DR SMART; SMO0209; TSP1.1.
DR PROSITE; PS50092; TSP1.1.
DR Malaria; Sporozoite; Repeat; Signal.
FT SIGNAL 1 16 PROBABLE.
FT CHAIN 17 388 CIRCUMSPOROZOITE PROTEIN.
FT DOMAIN 120 267 62 X 4 AA TANDEM REPEATS OF N-A-N-P.
FT DOMAIN 313 366 TSP TYPE-1.
SQ SEQUENCE 388 AA; 42245 MW; C01BEFBEZ35604 CRC64;

Query Match. 100.0%; Score 32; DB 1; Length 388;
Best Local Similarity 100.0%; Pred. No. 6.1;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KIXOPG 6
Db 112 KIXOPG 117

RESULT 5
CSP_PLAFA STANDARD; PRT; 412 AA.
AC P02893;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Circumsporozoite protein precursor (CS).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84250215; PubMed=6204383;
RA Dane J.B., Williams J.L., McCutchan T.F., Weber J.L., Witz R.A.,
RA Hochmeyer W.T., Maloy W.L., Haynes J.D., Schneider I., Roberts D.,
RA Sanders G.S., Reddy B.P., Diggs C.L., Miller L.H.;
RT "Structure of the gene encoding the immunodominant surface antigen on
RT the sporozoite of the human malaria parasite Plasmodium falciparum."
RL Science 225:593-599(1984).
CC -1- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
CC SURFACE ANTIGEN ON THE SPOOROZOITE (THE INFECTIVE STAGE OF THE
CC MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
CC VERTEBRATE HOST).

```

```

CC -1- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
CC ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
CC WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -1- SIMILARITY: Contains 1 TSP type-1 domain.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; K02194; AAA29524.1; -.
DR PIR; A03388; OZZOAF.
DR InterPro; IPR003067; Circmsprzoite.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PRO1303; CRCMSPRZOITE.
DR SMART; SMO0209; TSP1.1.
DR PROSITE; PS50092; TSP1.1.
DR Malaria; Sporozoite; Repeat; Signal.
FT SIGNAL 1 16 PROBABLE.
FT CHAIN 17 412 CIRCUMSPOROZOITE PROTEIN.
FT DOMAIN 123 290 41 X 4 AA TANDEM REPEATS OF P-N-A-N.
FT DOMAIN 337 390 TSP TYPE-1.
SQ SEQUENCE 412 AA; 44420 MW; 1EEED3DE90965F8 CRC64;

Query Match. 100.0%; Score 32; DB 1; Length 412;
Best Local Similarity 100.0%; Pred. No. 6.6;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KIXOPG 6
Db 112 KIXOPG 117

RESULT 6
CSP_PLAFA STANDARD; PRT; 424 AA.
AC P13814;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Circumsporozoite protein precursor (CS).
OS Plasmodium falciparum (isolate t4 / Thailand).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5846;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87315205; PubMed=3306373;
RA del Portillo H.A., Nussenzweig R.S., Enea V.;
RT "Circumsporozoite gene of a Plasmodium falciparum strain from
RT Thailand."
RL Mol. Biochem. Parasitol. 24:289-294(1987).
CC -1- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
CC SURFACE ANTIGEN ON THE SPOOROZOITE (THE INFECTIVE STAGE OF THE
CC MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
CC VERTEBRATE HOST).
CC -1- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
CC ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
CC WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -1- SIMILARITY: Contains 1 TSP type-1 domain.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; M19752; AAA29555.1; -.

```

```

DR PIR; A54533; A54533.
DR InterPro; IPR003067; Crcmsprzoite.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00090; TSP_1; 1.
DR PRINTS; PR01303; Crcmsprzoite.
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PS50092; TSP1; 1.
DR Malaria; Sporozoite; Repeat; Signal.
FT SIGNAL 1 16 CIRCUMSPOROZOITE PROTEIN.
FT CHAIN 17 424 PROBABLE.
FT DOMAIN 123 300 45 X 4 AA TANDEM REPEATS OF N-A-N-P.
FT DOMAIN 349 402 TSP TYPE-1.
SQ SEQUENCE 424 AA; 45610 MW; 710AB14238786CD9 CRC64;

Query Match 100.0%; Score 32; DB 1; Length 424;
Best Local Similarity 100.0%; Pred. No. 6.8;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLKQPG 6
Db 112 KLKQPG 117

RESULT 7
CSP_PIAFW STANDARD; PRT; 442 AA.
ID CSP_PIAFW STANDARD; PRT; 442 AA.
AC P08307;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Circumsporozoite protein precursor (CS).
OS Plasmodium falciparum (Isolate Wellcome).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5848;
RN [1]
RX MEDLINE=8715616; PubMed=3543671;
RA Lockyer M.J., Schwarz R.T.;
RT "Strain variation in the circumsporozoite protein gene of Plasmodium falciparum."
RL Mol. Biochem. Parasitol. 22:101-108(1987).
CC -1- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT SURFACE ANTIGEN ON THE SPOOROZOITE (THE INFECTIVE STAGE OF THE MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE VERTEBRATE HOST).
CC -1- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -1- SIMILARITY: Contains 1 TSP type-1 domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch).
CC -----
DR EMBL; M15505; AAA29554.1; -.
DR PIR; A54529; A54529.
DR InterPro; IPR003067; Crcmsprzoite.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00090; TSP_1; 1.
DR PRINTS; PR01303; Crcmsprzoite.
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PS50092; TSP1; 1.
DR Malaria; Sporozoite; Repeat; Signal.
FT SIGNAL 1 16 CIRCUMSPOROZOITE PROTEIN
FT CHAIN 17 424 47 X 4 AA TANDEM REPEATS OF N-A-N-P.
FT DOMAIN 130 320 TSP TYPE-1.
FT DOMAIN 367 420 TSP TYPE-1.
SQ SEQUENCE 442 AA; 47402 MW; BD57A9A152B85E03 CRC64;

```

```

Query Match 100.0%; Score 32; DB 1; Length 442;
Best Local Similarity 100.0%; Pred. No. 7.1;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLKQPG 6
Db 122 KLKQPG 127

RESULT 8
PKCB_HUMAN STANDARD; PRT; 1186 AA.
ID PKCB_HUMAN STANDARD; PRT; 1186 AA.
AC Q9ULU4; Q13117; Q8WXC5; Q9HIF3; Q9HIF4; Q9HIF5; Q9HIL8; Q9HIL9;
AC Q9H2G5; Q9NYN3; Q9UIX6;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Protein Kinase C binding protein 1 (Rack7) (Cutaneous T-cell lymphoma associated antigen se14-3) (CTCL tumor antigen se14-3) (Zinc finger MYND domain containing protein 8).
OS Homo sapiens (Human)
GN PRKCBP1 OR RACK7 OR ZMYND8 OR KIAA1125.
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RX MEDLINE=20458872; PubMed=11003709;
RA Posey S.C., Kuroda S., Price J.A., Pendleton J.K., Freedman B.I., Bowden D.W.;
RT "Identification and characterization of PRKCBP1, a candidate RACK-like protein."
RL Mamm. Genome 11:919-925(2000).
RN [2]
RX MEDLINE=20039618; PubMed=10574461;
RA Hirosewa M., Nagase T., Ishikawa K.-I., Kikuno R., Nomura N., Ohara O.;
RT "Characterization of cDNA clones selected by the Genemark analysis from size-fractionated cDNA libraries from human brain."
RL DNA Res. 6:329-336(1999).
RN [3]
RX MEDLINE=20039618; PubMed=10574461;
RA Hirosewa M., Nagase T., Ishikawa K.-I., Kikuno R., Nomura N., Ohara O.;
RT "Characterization of cDNA clones selected by the Genemark analysis from size-fractionated cDNA libraries from human brain."
RL DNA Res. 6:329-336(1999).
RN [4]
RX MEDLINE=21638749; PubMed=11780052;
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R., Jones M., Stevrides G., Almeida J.P., Babage A.K., Bagunuly C.L., Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M., Beasley O.P., Bird C.P., Blakey S.E., Bridgman A.W., Brown A.J., Buck D., Burrill W., Butler A.P., Carder C., Carter N.P., Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M., Clegg S., Cobley V.E., Collier R.E., Connor R., Corby N.R., Coulson A., Coville G.J., Deadman R., Dhami P., Dunn M., Ellington A.G., Frankland J.A., Fraser A., French L., Garner P., Graffham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E., Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J., Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D., Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S., Levaselaho M.H., Leversha M., Lloyd C., Lloyd D.M., Lovell J.D., Marsh V.L., Martin S.L., McCormachie L.J., McEay K., McMurtry A.A., Milne S., Mistry D., Moore M.-J.F., Mullikin J.C., Nickerson T., Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I., Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,

```



```

DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Indole-3-glycerol phosphate synthase (EC 4.1.1.48) (IGPS).
OS TRPC.
OC Lactobacillus casei.
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxID=1582;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90299861; PubMed=2113923;
RA Natori Y., Kano Y., Imamoto F.;
RT "Nucleotide sequences and genomic constitution of five tryptophan
RT genes of Lactobacillus casei."
RL J. Biochem. 107:248-255(1990).
CC -1- CATALYTIC ACTIVITY: 1-(2-carboxyphenylamino)-1-deoxy-D-ribose 5-
CC phosphate = 1-(indol-3-yl)glycerol 3-phosphate + CO(2) + H(2)O.
CC -1- PATHWAY: Tryptophan biosynthesis; fourth step.
CC -1- SIMILARITY: BELONGS TO THE TRPC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D00496; BAA00384.1; -.
DR PIR; S42344; JS0341.
DR HSSP; P00909; 1P1I.
DR HAMAP; MF_00134; -. 1.
DR InterPro; IPR003009; FMN enzyme.
DR InterPro; IPR001468; IGPS.
DR Pfam; PF00218; IGPS; 1.
DR ProDom; PD01511; IGPS; 1.
DR PROSITE; PS00614; IGPS; 1.
KM Tryptophan biosynthesis; Lyase; Decarboxylase.
SQ SEQUENCE 260 AA; 28943 MW; 47F572275A1BE8B CRC64;

Query Match 90.6%; Score 29; DB 1; Length 260;
Best Local Similarity 83.3%; Pred. No. 20;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLRPG 6
Db 44 KLRPG 49

RESULT 10
DHAA_AGR5 STANDARD; PRT; 304 AA.
AC Q8U671;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Haloalkane dehalogenase (EC 3.8.1.5).
GN DHAA OR DHA OR ATU6064 OR AGR_PTI_130.
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Plasmid pTIC58.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
OX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608550; PubMed=11743193;
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F., Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eissen J.A., Karp P.D., Bovee D., St.,
RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA Kutayavin T., Levy R., Li M.-Y., McClelland E., Palmieri A.,
RA Raymond C., Rouse G., Saenphimachak C., Wu Z., Romero P., Gordon D.,

```

```

RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Neeter E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
RT C58."
RL Science 294:2317-2323(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608551; PubMed=11743194;
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA Qurollo B., Goldman B.S., Cao Y., Akenazi M., Halling C., Mullin L.,
RA Houmel K., Gordon J., Vaudin M., Tatchouk O., Epp A., Liu F.,
RA Woliam C., Allinger M., Doughly D., Scott C., Lappas C., Markelz B.,
RA Planagan C., Crowell C., Gursan J., Lomo C., Sear C., Strub G.,
RA Cielo C., Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent
RT Agrobacterium tumefaciens C58."
RL Science 294:2323-2328(2001).
CC -1- FUNCTION: Catalyzes hydrolytic cleavage of carbon-halogen bonds in
CC halogenated aliphatic compounds, leading to the formation of the
CC corresponding primary alcohols, halide ions and protons (By
CC similarity).
CC -1- CATALYTIC ACTIVITY: 1-haloalkane + H(2)O = a primary alcohol +
CC halide.
CC -1- SUBUNIT: Monomer (By similarity).
CC -1- SIMILARITY: Belongs to the haloalkane dehalogenase family.
CC Subfamily 2.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE009425; AAL46300.1; ALT_INIT.
DR EMBL; AE007933; AAK91024.1; -.
DR HAMAP; MF_01231; -. 1.
DR InterPro; IPR000073; A/b_hydrolase.
DR InterPro; IPR000639; Epox_hydrolase.
DR InterPro; IPR000379; Ser_estrs_site.
DR Pfam; PF00561; abhydrolase; 1.
DR PRINTS; PR00412; EPOXYDRASE.
KM Hydrolyase; Plasmid; Complete proteome.
FT ACT_SITE 114 114 NUCLEOPHILE (BY SIMILARITY).
FT ACT_SITE 138 138 ACID (BY SIMILARITY).
FT ACT_SITE 280 280 BASE (BY SIMILARITY).
SQ SEQUENCE 304 AA; 34390 MW; 4783D160D4C86DFA CRC64;

Query Match 90.6%; Score 29; DB 1; Length 304;
Best Local Similarity 83.3%; Pred. No. 24;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLRPG 6
Db 159 KLRPG 164

RESULT 11
EPB3_HUMAN STANDARD; PRT; 998 AA.
AC P54753;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Epurin type-B receptor 3 precursor (EC 2.7.1.112) (Tyrosine-protein
DE kinase receptor HEK-2).
GN EPB3 OR ETR2 OR HEK2.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

```


[3]
 RN SEQUENCE OF 105-1083 FROM N.A., AND PARTIAL SEQUENCE.
 RP MEDLINE=97098442; PubMed=8942982;
 RA Tanese N., Saluja D., Vassallo M.F., Chen J.-L., Admon A.;
 RT "Molecular cloning and analysis of two subunits of the human TFIID
 RL complex: hTAF110 and hTAF1100.";
 CC PLoS Biol. 1:1361-1366 (1996).
 CC -1- FUNCTION: MAKES PART OF TFIID IS A MULTIMERIC PROTEIN COMPLEX THAT
 CC PLAYS A CENTRAL ROLE IN MEDIATING PROMOTER RESPONSES TO VARIOUS
 CC ACTIVATORS AND REPRESSORS. POTENTIATES TRANSCRIPTIONAL ACTIVATION
 CC BY THE AP-2 OF THE RETINOIC ACID, VITAMIN D3 AND THYROID HORMONE.
 CC -1- SUBUNIT: TFIID IS COMPOSED OF TATA BINDING PROTEIN (TBP) AND A
 CC NUMBER OF TBP-ASSOCIATED FACTORS (TAFs).
 CC -1- SIMILARITY: BELONGS TO THE TAF2C FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: Y11354; CAA72189.1; -;
 CC EMBL: AL137077; CAC36006.1; -;
 CC EMBL: AL109911; CAC22312.2; -;
 CC EMBL: U75308; AAC50901.1; -;
 CC PDB: 1H30; 26-SEP-02.
 CC TRANSEFACT; T02328; -;
 CC GeneW; HGNC:11537; TAF4.
 CC MIM: 601796; -;
 CC GO: GO:0005669; C:transcription factor TFIID complex; TAS.
 CC GO: GO:0016251; F:general RNA polymerase II transcription fac. . .; TAS.
 CC GO: GO:0005515; F:protein binding activity; TAS.
 CC GO: GO:0003713; F:transcription co-activator activity; TAS.
 CC InterPro: IPR003894; TAF_hom.
 CC Pfam: PF05236; TAF4; 1.
 CC SMART: SMO0549; TAFH; 1.
 CC KW Transcription regulation; Nuclear protein; 3D-structure.
 FT DOMAIN 39 42 POLY-HIS.
 FT 52 57 POLY-ALA.
 FT DOMAIN 98 101 POLY-GLY.
 FT DOMAIN 142 148 POLY-ALA.
 FT DOMAIN 268 275 POLY-PRO.
 FT DOMAIN 331 337 POLY-ALA.
 FT DOMAIN 680 683 POLY-PRO.
 FT DOMAIN 808 813 POLY-ALA.
 FT DOMAIN 828 831 POLY-ASP.
 FT CONFLICT 105 117 PGPPRRRLVPA -> GRGLQGRGREGS
 FT (IN REF. 3).
 FT CONFLICT 136 136 A -> S (IN REF. 2).
 FT CONFLICT 185 185 G -> GPG (IN REF. 2).
 FT CONFLICT 233 264 MISSING (IN REF. 3).
 FT CONFLICT 293 293 P -> L (IN REF. 3).
 SQ SEQUENCE 1083 AA; 109943 MW; A645382757A0752 CRC64;
 Query Match 90.6%; Score 29; DB 1; Length 1083;
 Best local Similarity 83.3%; Pred. No. 1e+02; Indels 0; Gaps 0;
 Matches 5; Conservative 1; Mismatches 0;

DE [Trypophanyl-tRNA synthetase, mitochondrial (EC 6.1.1.2) (Trypophan--
 DE tRNA ligase) (TPRS).
 GN MSW1 OR MSW OR YDR268W OR D9954.7.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 ON NCBI_TaxId=4932;
 RX [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86059396; PubMed=2999114;
 RA Myers A.M., Tzagoloff A.;
 RT "MSW, a yeast gene coding for mitochondrial tryptophanyl-tRNA
 RL synthetase.";
 RL J. Biol. Chem. 260:15371-15377 (1985).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92405222; PubMed=1525855;
 RA Entrop R., Langgut W., Lisowsky T., Schweizer E.;
 RT "An yeast nuclear mutation conferring temperature-sensitivity to the
 RL mitochondrial tryptophanyl-tRNA synthetase.";
 RN Curr. Genet. 21:281-283 (1992).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S28c / AB972;
 RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z.,
 RA Vassallo A., Fulton L., Gattung S., Greco T., Kisten J.,
 RA Kucaba T., Hallsworth K., Hawkins J., Hillier L., Jier M.,
 RA Johnson D., Johnston L., Langston Y., Latreille P., Le T.,
 RA Mardis E., Meneses S., Miller N., Nhan M., Pauley A., Peluso D.,
 RA Riken L., Riles L., Taich A., Trevasakis B., Vignati D.,
 RA Wilcox L., Wohldman P., Vaudin M., Wilson R., Waterston R.;
 RA Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA (Trp) = AMP +
 CC phosphate + L-tryptophanyl-tRNA (Trp).
 CC -1- SUBUNIT: Homodimer.
 CC -1- SUBCELLULAR LOCATION: Mitochondrial matrix.
 CC -1- SIMILARITY: Belongs to class-I aminoacyl-tRNA synthetase family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M12081; AAA34809.1; -;
 CC EMBL: X66165; CAA46947.1; -;
 CC EMBL: U51030; AAB64452.1; -;
 CC PIR: S70128; YMBYM.
 CC HSSP: P00953; 1D2R.
 CC SGD: S0002676; MSW1.
 CC GO: GO:0004830; F:tryptophan-tRNA ligase activity; IMP.
 CC GO: GO:0006436; F:tryptophanyl-tRNA aminoacylation; IMP.
 CC InterPro: IPR002305; tRNA-synt_1b.
 CC InterPro: IPR004142; tRNA-synt_1.
 CC InterPro: IPR002306; Trp-tRNA-synt_1b.
 CC Pfam: PF00579; tRNA-synt_1b; 1.
 CC PRINTS: PR01039; TRNASYNTTRP.
 CC TIGRFAMs: TIGR00233; trps; 1.
 CC PROSITE: PS00178; AA_TRNA_LIGASE_1; 1.
 KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
 KW Mitochondrion.
 FT SITE 43 51 "HIGH" REGION.
 FT SITE 244 248 "KMSKS" REGION.
 FT BINDING 247 247 ATP (BY SIMILARITY).
 FT CONFLICT 268 271 KIRK -> RLE (IN REF. 1).
 FT CONFLICT 371 379 ADIHKINGF -> PTFIK (IN REF. 1).
 SQ SEQUENCE 379 AA; 43015 MW; 807B6D21A991B108 CRC64;
 Query Match 87.5%; Score 28; DB 1; Length 379;
 Best local Similarity 83.3%; Pred. No. 54;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KIKOPG 6
:|||||
DB 64 ELKOPG 69

RESULT 14

HDA9_MOUSE STANDARD; PRT; 588 AA.
AC 099N13; Q9EP72;
DT 16-OCT-2001 (Rel. 40, Last Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Histone deacetylase 9 (HD9) (Histone deacetylase-related
protein) (MEP2-interacting transcription repressor MTR).
GN HDAC9 OR HDAC7B OR HDRP OR MTR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Swiss Webster / NIH;
RA Zhou X., Richon V.M., Rifkind R.A., Marks P.A.;
RT "Cloning of the mouse HDRP cDNA."
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=NIH Swiss; TISSUE=Heart;
RX PubMed=11022042;
RA Zhang C.L., McKiney T.A., Lu J.R., Olson E.N.;
RT "Association of COOH-terminal-binding protein (CtBP) and
MEP2-interacting transcription repressor (MTR) contributes to
transcriptional repression of the MEP2 transcription factor."
RL J. Biol. Chem. 276:35-39(2001).

- FUNCTION: Responsible for the deacetylation of lysine residues on
the N-terminal part of the core histones (H2A, H2B, H3 and H4).
Histone deacetylation gives a tag for epigenetic repression and
plays an important role in transcriptional regulation, cell cycle
progression and developmental events. Histone deacetylases act via
the formation of large multiprotein complexes.
- SUBUNIT: Interacts with MEP2.
- SIMILARITY: BELONGS TO THE HISTONE DEACETYLASE / ACUC / APHA
FAMILY. HD SUBFAMILY 2.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; AF235053; AAK15027.1; -
DR EMBL; AF324492; AAG48332.1; -
DR MGD; MGI:1931221; Hdac9.
DR GO; GO:0005634; C:nucleus, IDA.
DR GO; GO:0005515; F:protein binding activity, IPI.
DR GO; GO:0003714; F:transcription co-repressor activity, IDA.
DR GO; GO:0016564; F:transcriptional repressor activity, IDA.
DR GO; GO:0000122; P:negative regulation of transcription from P...; IDA.
DR GO; GO:0016481; P:negative regulation of transcription, IDA.
KW Hydroxylase, Nuclear protein, Chromatin regulator;
KW Transcription regulation, Repressor.
FT CONFLICT 120 120 K -> R (IN REF. 2).
FT CONFLICT 136 136 K -> R (IN REF. 2).
FT CONFLICT 177 178 MISSING (IN REF. 2).
FT CONFLICT 523 523 N -> T (IN REF. 2).
SQ SEQUENCE 588 AA, 65630 MW, D4BDF3774B3D99EA CRC64;
Query Match 87.5%; Score 28; DB 1; Length 588;
Best Local Similarity 83.3%; Pred. No. 87;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 KIKOPG 6
:|||||
DB 493 ELKOPG 498

RESULT 15

XE7_HUMAN STANDARD; PRT; 695 AA.
AC 002840; Q02832;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE B-lymphocyte antigen precursor (B-lymphocyte surface antigen) (721P)
DE (Protein XE7).
GN (XE7X OR XE7 OR DXYS155E) AND (XE7Y OR XE7 OR DXYS155E).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM LONG).
RX MEDLINE=93258310; PubMed=1302606;
RA Ellison J.W., Ramos C., Yen P.H., Shapiro L.J.;
RT "Structure and expression of the human pseudautosomal gene XE7."
RL Hum. Mol. Genet. 1:691-696(1992).
[2]
RP SEQUENCE FROM N.A. (ISOFORM SHORT).
RX MEDLINE=93066251; PubMed=1438229;
RA Voland J.R., Wyzkowski R.J., Huang M., Dutton R.W.;
RT "Cloning and sequencing of a trophoblast-endothelial-activated
lymphocyte surface protein: cDNA sequence and genomic structure."
RL Proc. Natl. Acad. Sci. U.S.A. 89:10425-10429(1992).

- FUNCTION: THIS PROTEIN MAY BE INVOLVED IN B-CELL ACTIVATION. MAY
ALSO BE INVOLVED IN SIGNAL TRANSDUCTION AND GENE REGULATION.
- SUBUNIT: Monomer.
- SIMILARITY: BELONGS TO THE HISTONE DEACETYLASE / ACUC / APHA
FAMILY. HD SUBFAMILY 2.
- ALTERNATIVE PRODUCTS: Membrane-associated (Probable).
Event=Alternative splicing; Named isoforms=2;
Name=Long;
Name=Short;
IsoId=002040-1; Sequence=Displayed;
IsoId=002040-2; Sequence=VSP 004490, VSP 004491.

CC TISSUE SPECIFICITY: EXPRESSED IN ACTIVATED LYMPHOCYTES, MOST
CC VASCULAR ENDOTHELIAL, AND SYNCYTIOTROPHOBLAST.
CC - CAUTION: REF.2 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A
CC FRAME SHIFT IN POSITION 461.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; L03426; AAA61304.1; -
DR EMBL; L03426; AAA61303.1; -
DR EMBL; M99578; AAA36187.1; ALT_FRAME.
DR PIR; A46419; A46419.
DR PIR; A54325; A54325.
DR MIM; 312095; -
DR MIM; 465000; -
DR GO; GO:0042113; P:B-cell activation, NAS.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; NAS.
DR GO; GO:0007165; P:signal transduction, NAS.
KW Glycoprotein, Signal, Membrane, Alternative splicing.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 695 B-LYMPHOCYTE ANTIGEN.
FT CARBOHYD 85 85 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 138 138 N-LINKED (GLCNAC...) (POTENTIAL).

FT CARBOHYD 513 513 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 566 566 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 385 385 A -> L (in isoform Short).
FT VARSPLIC 385 385 /FtId=VSP_004490.
FT VARSPLIC 386 695 Missing (in isoform Short).
FT VARSPLIC 386 695 /FtId=VSP_004491.
FT CONFLICT 443 443 L -> Q (IN REF. 2).
FT CONFLICT 500 500 A -> P (IN REF. 2).
FT CONFLICT 502 502 H -> P (IN REF. 2).
SQ SEQUENCE 695 AA; 80709 NM; 9FE9D9DF617BFDCE8 CRC64;

Query Match 87.5%; Score 28; DB 1; Length 695;
Best Local Similarity 83.3%; Pred. No. 1.le+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKKOPG 6
DB 38 QKKOPG 43

Search completed: December 23, 2003, 15:33:44
Job time : 6 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compen Ltd.

OM protein - protein search, using sw model

Run on: December 23, 2003, 15:33:08 ; Search time 14 Seconds

(without alignments)
110.594 Million cell updates/sec

Title: US-09-980-564-12

Perfect score: 32

Sequence: 1 KLKOPG 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_23:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp Vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query Length	DB ID	Description
1	32	100.0	226	5 Q26136	Q26136 plasmodium
2	32	100.0	246	5 Q81SA7	Q81SA7 plasmodium
3	32	100.0	249	5 Q9BMN5	Q9BMN5 plasmodium
4	32	100.0	260	5 Q81SA8	Q81SA8 plasmodium
5	32	100.0	263	5 Q9GQPI	Q9GQPI plasmodium
6	32	100.0	277	5 Q9GQPO	Q9GQPO plasmodium
7	32	100.0	281	16 Q8Y1F6	Q8Y1F6 bruceella me
8	32	100.0	281	16 Q8ZFZ5	Q8ZFZ5 bruceella su
9	32	100.0	288	5 Q9GQMO	Q9GQMO plasmodium
10	32	100.0	288	5 Q9GQMI	Q9GQMI plasmodium
11	32	100.0	302	5 Q9GQI8	Q9GQI8 plasmodium
12	32	100.0	302	5 Q9GQI9	Q9GQI9 plasmodium
13	32	100.0	303	5 Q9BMN3	Q9BMN3 plasmodium
14	32	100.0	303	5 Q9BMN4	Q9BMN4 plasmodium
15	32	100.0	317	5 Q9BH54	Q9BH54 plasmodium
16	32	100.0	318	5 Q9BKJ7	Q9BKJ7 plasmodium

17	32	100.0	339	5 Q26120	Q26120 plasmodium
18	32	100.0	343	5 Q26119	Q26119 plasmodium
19	32	100.0	350	5 Q26118	Q26118 plasmodium
20	32	100.0	350	5 Q26117	Q26117 plasmodium
21	32	100.0	353	5 Q81TW6	Q81TW6 plasmodium
22	32	100.0	356	5 Q9BPK3	Q9BPK3 plasmodium
23	32	100.0	356	5 Q9BPK7	Q9BPK7 plasmodium
24	32	100.0	356	5 Q819G5	Q819G5 plasmodium
25	32	100.0	364	5 Q819G4	Q819G4 plasmodium
26	32	100.0	367	5 Q26169	Q26169 plasmodium
27	32	100.0	368	5 Q03753	Q03753 plasmodium
28	32	100.0	368	5 Q8MPK0	Q8MPK0 plasmodium
29	32	100.0	368	5 Q8MPJ9	Q8MPJ9 plasmodium
30	32	100.0	368	5 Q8MPK1	Q8MPK1 plasmodium
31	32	100.0	368	5 Q819G3	Q819G3 plasmodium
32	32	100.0	370	5 Q9BPK6	Q9BPK6 plasmodium
33	32	100.0	370	5 Q8MPK2	Q8MPK2 plasmodium
34	32	100.0	370	5 Q9BPK8	Q9BPK8 plasmodium
35	32	100.0	377	5 Q962K4	Q962K4 plasmodium
36	32	100.0	377	5 Q26116	Q26116 plasmodium
37	32	100.0	380	5 Q819G8	Q819G8 plasmodium
38	32	100.0	380	5 Q819G1	Q819G1 plasmodium
39	32	100.0	383	5 Q9GPN1	Q9GPN1 plasmodium
40	32	100.0	389	5 Q819J2	Q819J2 plasmodium
41	32	100.0	392	5 Q819G2	Q819G2 plasmodium
42	32	100.0	393	5 Q8MPK3	Q8MPK3 plasmodium
43	32	100.0	393	5 Q819J1	Q819J1 plasmodium
44	32	100.0	393	5 Q819J0	Q819J0 plasmodium
45	32	100.0	396	5 Q03136	Q03136 plasmodium

ALIGNMENTS

RESULT 1
Q26136 PRELIMINARY; PRT; 226 AA.
AC Q26136;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE CSP protein (Fragment).
OS CSP.
GN Plasmodium vivax.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_taxid=5855;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95043163; PubMed=7955170;
RA Huang T., Cheng O., Allan S., Huang Y.;
RT "DNA sequencing of circumsporozoite protein genes of Plasmodium vivax
from four different countries.";
RT study on the flank sequences]."
RL Chung Kuo Chi Sheng Chung Hsueh Yu Chi Sheng Chung 12:85-92(1994).
DR EMBL; S73385; A014117.1; -
DR InterPro: IPR000884; TSP1.
DR Pfam; PF00090; CSP_1; 1.
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PS50092; TSP1; 1.
DR NON TER
SQ SEQUENCE 226 AA; 23156 MW; A56FFCC4082E018F CRC64;
Query Match 100.0%; Score 32; DB 5; Length 226;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLKOPG 6
|||
Db 68 KLKOPG 73

RESULT 2
Q81SA7

```

ID 08ISA7 PRELIMINARY; PRT; 246 AA.
AC 08ISA7;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Circumsporozoite protein (Fragment).
GN CSP.
OS Plasmodium vivax.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5855;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PvSK2;
RA Kim T.-Y., Kang S.-Y., Kim T.-S., Chai J.-Y., Lee S.-H., Hong S.-J.;
RT "Circumsporozoite protein (CSP) PvSK2 from Plasmodium vivax South
RT Korean isolate.";
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY137778; AAN34503.1; -.
FT NON_TER 1
FT NON_TER 246
SQ SEQUENCE 246 AA; 22926 MW; 23B2B619FAFDCl0 CRC64;

Query Match 100.0%; Score 32; DB 5; Length 246;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 KLKOPG 6
   |||||
Db 28 KLKOPG 33

RESULT 3
ID 09BMN5 PRELIMINARY; PRT; 249 AA.
AC 09BMN5;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Circumsporozoite protein (Fragment).
GN Plasmodium vivax.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5855;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KNH CSP99-069;
RA Lee H., Cho S., Kim T.;
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF316580; AAG60034.1; -.
DR InterPro; IPR003067; Crcmsprzoite.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00090; tsp.1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PS50092; TSP1; 1.
FT NON_TER 1
FT NON_TER 249
SQ SEQUENCE 249 AA; 24256 MW; 224A1F6ADB9FD013 CRC64;

```

```

Query Match 100.0%; Score 32; DB 5; Length 249;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLKOPG 6
   |||||
Db 15 KLKOPG 20

```

```

RESULT 4
ID 08ISA8 PRELIMINARY; PRT; 260 AA.
AC 08ISA8;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)

```

```

DE Circumsporozoite protein (Fragment).
GN CSP.
OS Plasmodium vivax.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5855;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PvSK1;
RA Kim T.-Y., Kang S.-Y., Kim T.-S., Chai J.-Y., Lee S.-H., Hong S.-J.;
RT "Circumsporozoite protein (CSP) PvSK1 from Plasmodium vivax South
RT Korean isolate.";
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY137777; AAN34502.1; -.
FT NON_TER 1
FT NON_TER 260
SQ SEQUENCE 260 AA; 23583 MW; 590BF7023CD160EA CRC64;

Query Match 100.0%; Score 32; DB 5; Length 260;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLKOPG 6
   |||||
Db 28 KLKOPG 33

```

```

RESULT 5
ID 09GQP1 PRELIMINARY; PRT; 263 AA.
AC 09GQP1;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Circumsporozoite protein (Fragment).
GN CSP.
OS Plasmodium vivax.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5855;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SK-A;
RX MEDLINE=20099860; PubMed=10634043;
RA Kim W.-G., Park Y.H., Chung J.-Y., Kim J.-P., Hong S.-T., Lee W.-J.,
RA Kim T.-S., Lee J.-S.;
RT "Two new genotypes of Plasmodium vivax circumsporozoite protein found
RT in the Republic of Korea.";
RL Korean J. Parasitol. 37:265-270(1999).
DR EMBL; AF215740; AAG43992.1; -.
DR InterPro; IPR003067; Crcmsprzoite.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00090; tsp.1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
FT NON_TER 1
FT NON_TER 263
SQ SEQUENCE 263 AA; 24671 MW; 822BD283DEBA9E7 CRC64;

```

```

Query Match 100.0%; Score 32; DB 5; Length 263;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLKOPG 6
   |||||
Db 26 KLKOPG 31

```

```

RESULT 6
ID 09GQP0 PRELIMINARY; PRT; 277 AA.
AC 09GQP0;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Circumsporozoite protein (Fragment).

```

GN CSP
 OS Plasmodium vivax.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5855;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SK-B;
 RX MEDLINE=20099860; PubMed=10634043;
 RA Kho W.G., Park Y.H., Chung J.Y., Kim J.P., Hong S.T., Lee W.J.,
 Kim T.S., Lee J.S.;
 RT "Two new genotypes of Plasmodium vivax circumsporozoite protein found
 in the Republic of Korea";
 RL Korean J. Parasitol. 37:265-270(1999).
 DR EMBL; AF215741; AAG4393.1; -;
 DR InterPro; IPR003067; Circsprzoite.
 DR Pfam; PF000090; tsp.1; 1.
 DR PRINTS; PR01303; CRCMSPRZOITE.
 FT NON_TER 1
 FT NON_TER 1
 SQ SEQUENCE 277 AA; 25328 MW; B2BDAC4BDF72C374 CRC64;

Query Match 100.0%; Score 32; DB 5; Length 277;
 Best Local Similarity 100.0%; Pred. No. 28;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KUKOPG 6
 |||||
 DB 26 KUKOPG 31

RESULT 7
 ID O8YIF6 PRELIMINARY; PRT; 281 AA.
 AC O8YIF6;

DT 01-MAR-2002 (TREMBLrel. 20, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Polipoprotein diacylglycerol transferase (EC 2.4.99.-).
 GN BME10488.

OS Brucella melitensis.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Brucellaceae; Brucella.
 OX NCBI_TaxID=29459;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=16M / ATCC 23456 / Biotype 1;
 RX MEDLINE=20020109; PubMed=11756688;
 RA DelVecchio V.G., Kapatali V., Redkar R.J., Patra G., Mufar C., Los T.,
 Ivanova N., Anderson I., Bhattacharya A., Lykidis A., Reznik G.,
 Jablonaki L., Larsen N., D'Souza M., Bernal A., Mazur M., Goldsman E.,
 Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,
 Haeflorn R., Kyprides N., Overbeek R.;
 RT "The genome sequence of the facultative intracellular pathogen
 Brucella melitensis";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).

DR EMBL; AE009491; AAL51669.1; -;
 DR InterPro; IPR00140; LCT.
 DR Pfam; PF01790; LCT; 1.
 DR ProDom; PD005412; LCT; 1.
 DR TIGRFAMs; TIGR00544; LCT; 1.
 KW Transferase; Glycosyltransferase; Complete proteome.
 SQ SEQUENCE 281 AA; 31000 MW; 230D15039F6C1AB1 CRC64;

Query Match 100.0%; Score 32; DB 16; Length 281;
 Best Local Similarity 100.0%; Pred. No. 29;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KUKOPG 6
 |||||
 DB 212 KUKOPG 217

RESULT 8
 ID O8FZF5 PRELIMINARY; PRT; 281 AA.
 AC O8FZF5;
 DT 01-MAR-2003 (TREMBLrel. 23, Created)
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Polipoprotein diacylglycerol transferase.
 GN LCT OR BR1528.
 OS Brucella suis.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Brucellaceae; Brucella.
 OX NCBI_TaxID=29461;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1330 / Biovar 1;
 RX MEDLINE=22247741; PubMed=12271122.
 RA Paulsen I.T., Seshadri R., Nelson K.E., Eisen J.A., Heidelberg J.F.,
 Read T.D., Dodson R.J., Umeyam L., Brinkac L.M., Beaman M.J.,
 Daugherty S.C., Debey R.T., Durkin A.S., Kolonay J.F., Madupu R.,
 Nelson W.C., Ayodeji B., Kraul M., Shetty J., Malek J., Van Aken S.E.,
 Riedmuller S., Tettelin H., Gill S.R., White O., Salzberg S.L.,
 Hoover D.L., Lindler L.E., Hailing S.M., Boyle S.M., Fraser C.M.;
 RT "The Brucella suis genome reveals fundamental similarities between
 animal and plant pathogens and symbionts";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:13148-13153(2002).
 DR EMBL; AE014447; AAN30438.1; -;
 DR TIGR; BR1528;
 KW Lipoprotein; Transferase; Complete proteome.
 SQ SEQUENCE 281 AA; 31000 MW; 230D15039F6C1AB1 CRC64;

Query Match 100.0%; Score 32; DB 16; Length 281;
 Best Local Similarity 100.0%; Pred. No. 29;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KUKOPG 6
 |||||
 DB 212 KUKOPG 217

RESULT 9

ID O9GQMO PRELIMINARY; PRT; 288 AA.

AC O9GQMO;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Circumsporozoite protein (Fragment).
 GN CSP.

OS Plasmodium vivax.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5855;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=SK98088;
 RA Kim M.-J., Lee S.-H.;
 RT "Plasmodium vivax circumsporozoite protein gene";
 RL Submitted (FE8-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF236847; AAG44386.1; -;
 DR InterPro; IPR003067; Circsprzoite.
 DR InterPro; IPR000884; TSP1.

DR Pfam; PF000090; tsp.1; 1.
 DR PRINTS; PR01303; CRCMSPRZOITE.
 DR SMART; SM00209; TSP1; 1.
 DR PROSITE; PS50092; TSP1; 1.
 FT NON_TER 1
 FT NON_TER 1
 SQ SEQUENCE 288 AA; 27232 MW; FF2FA15A6A16B2A2 CRC64;

Query Match 100.0%; Score 32; DB 5; Length 288;
 Best Local Similarity 100.0%; Pred. No. 29;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLKOPG 6
 |||||
 Db 5 KLKOPG 10

RESULT 10

Q9GOM1 PRELIMINARY; PRT; 288 AA.

AC Q9GOM1 (TREMblrel. 16, Created)
 DT 01-MAR-2001 (TREMblrel. 16, last sequence update)
 DT 01-MAR-2001 (TREMblrel. 22, last annotation update)
 DE Circumsporozite protein (Fragment).
 GN CSP.
 OS Plasmodium vivax.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 CX NCBI_TaxID=5855;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SK98001;
 RA Kim M.-J., Lee S.-H.;
 RT "Plasmodium vivax circumsporozite protein gene."
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF236846; AAC44385.1; -;
 DR InterPro; IPR003067; Crcmsprzoite.
 DR Pfam; PF00090; tsp_1; 1.
 DR PRINTS; PR01303; Crcmsprzoite.
 DR SMART; SM00209; TSP1; 1.
 DR PROSITE; PSS0092; TSP1; 1.
 FT NON_TER 1
 FT NON_TER 288
 SQ SEQUENCE 288 AA; 27319 MW; FF35159568749372 CRC64;

Query Match 100.0%; Score 32; DB 5; Length 288;
 Best Local Similarity 100.0%; Pred. No. 29;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLKOPG 6
 |||||
 Db 5 KLKOPG 10

RESULT 11

Q9GOL8 PRELIMINARY; PRT; 302 AA.

AC Q9GOL8 (TREMblrel. 16, Created)
 DT 01-MAR-2001 (TREMblrel. 16, last sequence update)
 DT 01-MAR-2001 (TREMblrel. 22, last annotation update)
 DE Circumsporozite protein (Fragment).
 GN CSP.
 OS Plasmodium vivax.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 CX NCBI_TaxID=5855;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SK974601;
 RA Kim M.-J., Lee S.-H.;
 RT "Plasmodium vivax circumsporozite protein gene."
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF236849; AAC44388.1; -;
 DR InterPro; IPR003067; Crcmsprzoite.
 DR InterPro; IPR000884; TSP1.
 DR Pfam; PF00090; tsp_1; 1.
 DR PRINTS; PR01303; Crcmsprzoite.
 DR SMART; SM00209; TSP1; 1.
 DR PROSITE; PSS0092; TSP1; 1.
 FT NON_TER 1
 FT NON_TER 302
 SQ SEQUENCE 302 AA; 28102 MW; 4A752DEF28A6224 CRC64;

Query Match 100.0%; Score 32; DB 5; Length 302;

Best Local Similarity 100.0%; Pred. No. 31;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLKOPG 6
 |||||
 Db 5 KLKOPG 10

RESULT 12

Q9GOL9 PRELIMINARY; PRT; 302 AA.

AC Q9GOL9 (TREMblrel. 16, Created)
 DT 01-MAR-2001 (TREMblrel. 16, last sequence update)
 DT 01-OCT-2002 (TREMblrel. 22, last annotation update)
 DE Circumsporozite protein (Fragment).
 GN CSP.
 OS Plasmodium vivax.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 CX NCBI_TaxID=5855;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SK98007;
 RA Kim M.-J., Lee S.-H.;
 RT "Plasmodium vivax circumsporozite protein gene."
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF236848; AAC44387.1; -;
 DR InterPro; IPR003067; Crcmsprzoite.
 DR InterPro; IPR000884; TSP1.
 DR Pfam; PF00090; tsp_1; 1.
 DR PRINTS; PR01303; Crcmsprzoite.
 DR SMART; SM00209; TSP1; 1.
 DR PROSITE; PSS0092; TSP1; 1.
 FT NON_TER 1
 FT NON_TER 302
 SQ SEQUENCE 302 AA; 28016 MW; E44EF30DD8C3ED7 CRC64;

Query Match 100.0%; Score 32; DB 5; Length 302;
 Best Local Similarity 100.0%; Pred. No. 31;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLKOPG 6
 |||||
 Db 5 KLKOPG 10

RESULT 13

Q9BMN3 PRELIMINARY; PRT; 303 AA.

AC Q9BMN3 (TREMblrel. 17, Created)
 DT 01-JUN-2001 (TREMblrel. 17, last sequence update)
 DT 01-JUN-2001 (TREMblrel. 22, last annotation update)
 DE Circumsporozite protein (Fragment).
 GN CSP.
 OS Plasmodium vivax.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 CX NCBI_TaxID=5855;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=KNH CSP99-099;
 RA Lee H., Cho S., Kim T.;
 RT Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF16582; AAC60056.1; -;
 DR InterPro; IPR003067; Crcmsprzoite.
 DR InterPro; IPR000884; TSP1.
 DR Pfam; PF00090; tsp_1; 1.
 DR PRINTS; PR01303; Crcmsprzoite.
 DR SMART; SM00209; TSP1; 1.
 DR PROSITE; PSS0092; TSP1; 1.
 FT NON_TER 1
 FT NON_TER 303
 SQ SEQUENCE 303 AA; 29002 MW; 7E70478AB0C6F94B CRC64;

Query Match 100.0%; Score 32; DB 5; Length 303;

Best Local Similarity 100.0%; Pred. No. 31;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KIKQPG 6
|||||
DB 15 KIKQPG 20

OY 1 KIKQPG 6
|||||
DB 15 KIKQPG 20

RESULT 14

O9BMN4 PRELIMINARY; PRT; 303 AA.
AC O9BMN4;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, last annotation update)
DE Circumsporozoite protein (Fragment).
OS Plasmodium vivax.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
NCBI_TaxID=5855;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=KN1H CSP-078;
RA Lee H., Cho S., Kim T.;
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF316581; AAG60055.1; -.
DR InterPro; IPR003067; Circmsprzoite.
DR InterPro; IPR000884; TSPL.
DR Pfam; PF00090; tspl.1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSPL.1.
DR PROSITE; PS50092; TSPL.1.
FT NON TER 1
SQ SEQUENCE .303 AA; 28944 MW; 5833DFE707550984 CRC64;

Query Match 100.0%; Score 32; DB 5; Length 303;

Best Local Similarity 100.0%; Pred. No. 31;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KIKQPG 6
|||||
DB 15 KIKQPG 20

RESULT 15

O9BH54 PRELIMINARY; PRT; 317 AA.
AC O9BH54;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, last annotation update)
DE Circumsporozoite protein (Fragment).
OS Plasmodium vivax.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
NCBI_TaxID=5855;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=KN1H CSP99-056, and KN1H CSP99-161;
RA Lee H., Cho S., Kim T.;
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF316579; AAG60053.1; -.
DR EMBL; AF316583; AAG60057.1; -.
DR InterPro; IPR003067; Circmsprzoite.
DR InterPro; IPR000884; TSPL.
DR Pfam; PF00090; tspl.1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSPL.1.
DR PROSITE; PS50092; TSPL.1.
FT NON TER 1
SQ SEQUENCE 317 AA; 29729 MW; B8102C6D9D005CF3 CRC64;

Query Match 100.0%; Score 32; DB 5; Length 317;

Best Local Similarity 100.0%; Pred. No. 32;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Search completed: December 23, 2003, 15:35:33
Job time : 15 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compen Ltd.

OM protein - protein search, using sw model

Run on: December 23, 2003, 15:33:08 ; Search time 53.25 Seconds
(without alignments)
53.654 Million cell updates/sec

Title: US-09-980-564-13

Perfect score: 103
Sequence: 1 EWSPTCTGNGIQYRIK 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A Geneseq_19Jun03:*

- 1: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT:*
- 2: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT:*
- 3: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT:*
- 4: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT:*
- 5: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT:*
- 6: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT:*
- 7: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT:*
- 8: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT:*
- 9: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT:*
- 10: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT:*
- 11: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT:*
- 12: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT:*
- 13: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT:*
- 14: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT:*
- 15: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT:*
- 16: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT:*
- 17: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT:*
- 18: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT:*
- 19: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT:*
- 20: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT:*
- 21: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT:*
- 22: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT:*
- 23: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT:*
- 24: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	103	100.0	18	AAW59270	P. falciparum circ
2	103	100.0	18	AAW59270	Peptide RII derive
3	103	100.0	23	AAW59271	P. falciparum circ
4	103	100.0	23	AAW59271	Peptide used for F
5	103	100.0	23	AAW59271	P. falciparum CSP
6	103	100.0	33	AAW59274	Peptide used for F
7	103	100.0	46	AAW59274	P. falciparum circ
8	103	100.0	54	AAW59274	Circumsporozoite R
9	103	100.0	126	AAW59273	P. falciparum circ

10	103	100.0	180	11	AAW07290	Circumsporozoite a
11	103	100.0	184	11	AAW07289	Circumsporozoite a
12	103	100.0	309	12	AAW13175	NS1 81-RLFdel1a9.
13	103	100.0	319	11	AAW07945	NS181RLFAuth plasm
14	103	100.0	319	12	AAW13176	NS1 81-RLFAuth. P
15	103	100.0	327	12	AAW13177	NS1 81-RLFAuth + (
16	103	100.0	335	12	AAW13178	NS1 81 (NANP) 4RLFAu
17	103	100.0	335	12	AAW13179	NS1 81 (NNDP) 4RLFAu
18	103	100.0	412	7	AAW60416	CS protein of mala
19	103	100.0	412	9	AAW80835	Sequence encoded b
20	103	100.0	424	14	AAW37796	RIS protein. Synt
21	103	100.0	424	14	AAW37797	RIS* protein. Syn
22	96	93.2	411	9	AAW83144	Sequence encoded b
23	86.5	84.0	21	15	AAW51427	Circumsporozoite p
24	83	80.6	18	15	AAW51429	Circumsporozoite p
25	83	80.6	18	15	AAW51432	Circumsporozoite p
26	83	80.6	19	15	AAW51436	Circumsporozoite p
27	83	80.6	20	16	AAW71646	CS Region II + mme
28	83	80.6	20	16	AAW71648	CS Region II + mme
29	83	80.6	20	20	AAW22138	P. falciparum CSP:
30	83	80.6	429	10	AAW90064	Antigenic protein
31	82	79.6	18	15	AAW51430	Circumsporozoite p
32	82	79.6	18	15	AAW51431	Circumsporozoite p
33	82	79.6	402	8	AAW70709	Plasmodium cynomol
34	80	77.7	16	15	AAW51428	Plasmodium cynomol
35	80	77.7	47	23	AAW82030	Hidden Markov mode
36	80	77.7	54	22	AAW848115	Consensus thrombo
37	80	77.7	54	23	AAW78139	Consensus sequenc
38	80	77.7	54	23	AAW19378	Human thrombospond
39	80	77.7	54	23	AAW07601	Thrombospondin typ
40	80	77.7	54	24	AAW08386	Consensus sequenc
41	80	77.7	2150	21	AAW53898	Amino acid sequenc
42	80	77.7	2165	22	AAW90617	Human secreted pro
43	77	74.8	20	14	AAW37985	P. vivax circumspo
44	77	74.8	53	22	AAW90621	Human secreted pro
45	77	74.8	134	22	AAW11292	Human HSP-124 pro

ALIGNMENTS

RESULT 1	AAW59270	AAW59270 standard; peptide; 18 AA.
ID	AAW59270;	
AC	AAW59270;	
XX		
DT	27-AUG-1998 (first entry)	
XX		
DE	P. falciparum circumsporozoite region II peptide fragment.	
XX		
KM	Malaria; gene therapy; hepatocyte; liver; circumsporozoite; ligand;	
KW	targeted delivery; therapy; disease; cancer; hepatitis; cystic fibrosis;	
KW	hypercholesterolaemia; phenylketonuria; haemophilia.	
XX		
OS	Plasmodium falciparum.	
XX		
FN	US5766899-A.	
XX		
PD	16-JUN-1998.	
XX		
PF	27-FEB-1995; 95US-0395602.	
XX		
PR	27-FEB-1995; 95US-0395602.	
XX		
PA	(TEXA) UNIV TEXAS SYSTEM.	
XX		
PI	Ding Z, Kuo MT;	
XX		
DR	WPI; 1998-361692/31.	
XX		
PT	Complexes for targeted delivery of nucleic acids to hepatocytes -	
PT	containing plasmodium circumsporozoite polypeptide as targeting	

PT ligand
 XX
 PS Claim 2; Column 29-30; 34pp; English.
 XX
 CC AAM59270-M59274 are fragments of malarial circumsporozoite (CS) region
 CC II isolated from Plasmodium falciparum. These fragments can be used as
 CC ligands in a method for the targeted delivery of nucleic acid to cells
 CC in culture or cells in vivo, especially where the cells are hepatocytes.
 CC Therapy of diseases such as cancer, malaria, hepatitis, cystic fibrosis,
 CC hypercholesterolaemia, phenylketonuria and haemophilia is mentioned.
 CC CS polypeptides are liver cell specific with rapid hepatic invasion.
 CC They are more efficient than the prior art asialoosomucoid (ASOR).
 CC ligands, of which there may be an accumulation in certain diseases due to
 CC receptor downregulation.
 XX
 SQ Sequence 18 AA;
 Query Match 100.0%; Score 103; DB 19; Length 18;
 Best Local Similarity 100.0%; Pred. No. 9,6e-08;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 EMSPCSVTCGNGIQVRIK 18
 1 EMSPCSVTCGNGIQVRIK 18
 Db 1 EMSPCSVTCGNGIQVRIK 18
 RESULT 2
 AAB49237
 ID AAB49237 standard; Protein; 18 AA.
 AC AAB49237;
 XX
 DT 14-MAR-2001 (first entry)
 XX
 DE Peptide RII derived from malaria circumsporozoite surface protein.
 XX
 KM Adenoviral vector; inverted terminal repeat; ITR; gene therapy;
 KM sickle cell anemia; thalassemia; atherosclerosis; restenosis;
 KM cancer.
 XX
 OS Synthetic.
 XX
 PN WO200073478-A2.
 PD 07-DEC-2000.
 PS 01-JUN-2000; 2000WO-US5442.
 PF 01-JUN-1999; 99US-0137213.
 PR 22-OCT-1999; 99US-0161097.
 XX
 PA (UNIV) UNIV WASHINGTON.
 XX
 PI Lieber A, Shayakhmetov D, Farrar D, Papayannopoulou T;
 XX
 DR WPI; 2001-049942/06.
 XX
 PT Recombinant adenoviral vector containing transgene, used to produce
 PT gutless vectors for gene therapy, targetable to selected cells and
 PT lacking antigenicity -
 XX
 PS Example 2; Page 97; 156pp; English.
 XX
 CC The present invention relates to a first generation recombinant
 CC adenoviral vector, part of which integrates into a host cell genome. Two
 CC inverted terminal repeats (ITRs) allow integration of a transgene into
 CC the host genome. The invention can be used to make 'gutless' vectors
 CC for gene therapy, e.g. of sickle cell anemia or thalassemia (targeting
 CC hematopoietic cells) or atherosclerosis or restenosis (targeting
 CC endothelial cells), or more generally a wide range of genetic diseases,
 CC cancers and infectious diseases.
 XX
 SQ Sequence 18 AA;

Query Match 100.0%; Score 103; DB 22; Length 18;
 Best Local Similarity 100.0%; Pred. No. 9,6e-08;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 EMSPCSVTCGNGIQVRIK 18
 1 EMSPCSVTCGNGIQVRIK 18
 Db 1 EMSPCSVTCGNGIQVRIK 18
 RESULT 3
 AAM59271
 ID AAM59271 standard; peptide; 23 AA.
 AC AAM59271;
 XX
 DT 27-AUG-1998 (first entry)
 XX
 DE P. falciparum circumsporozoite region II peptide fragment E35.
 XX
 KM Malaria; gene therapy; hepatocyte; liver; circumsporozoite; ligand;
 KM targeted delivery; therapy; disease; cancer; hepatitis; cystic fibrosis;
 KM hypercholesterolaemia; phenylketonuria; haemophilia.
 XX
 OS Plasmodium falciparum.
 XX
 PN US5766899-A.
 PD 16-JUN-1998.
 PS 27-FEB-1995; 95US-0395602.
 PF 27-FEB-1995; 95US-0395602.
 PR 27-FEB-1995; 95US-0395602.
 XX
 PA (TEXA) UNIV TEXAS SYSTEM.
 XX
 PI Ding Z, Kuo MT;
 XX
 DR WPI; 1998-361692/31.
 XX
 PT Complexes for targeted delivery of nucleic acids to hepatocytes -
 PT containing Plasmodium circumsporozoite polypeptide as targeting
 PT ligand
 XX
 PS Claim 2; Column 29-30; 34pp; English.
 XX
 CC AAM59270-M59274 are fragments of malarial circumsporozoite (CS) region
 CC II isolated from Plasmodium falciparum. These fragments can be used as
 CC ligands in a method for the targeted delivery of nucleic acid to cells
 CC in culture or cells in vivo, especially where the cells are hepatocytes.
 CC Therapy of diseases such as cancer, malaria, hepatitis, cystic fibrosis,
 CC hypercholesterolaemia, phenylketonuria and haemophilia is mentioned.
 CC CS polypeptides are liver cell specific with rapid hepatic invasion.
 CC They are more efficient than the prior art asialoosomucoid (ASOR).
 CC ligands, of which there may be an accumulation in certain diseases due to
 CC receptor downregulation.
 XX
 SQ Sequence 23 AA;
 Query Match 100.0%; Score 103; DB 19; Length 23;
 Best Local Similarity 100.0%; Pred. No. 1,2e-07;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 EMSPCSVTCGNGIQVRIK 18
 1 EMSPCSVTCGNGIQVRIK 18
 Db 1 EMSPCSVTCGNGIQVRIK 18
 RESULT 4
 AAY22153
 ID AAY22153 standard; Protein; 23 AA.
 AC AAY22153;

XX 08-SEP-1999 (first entry)
 DT Peptide used for FHV chimeric particle construction.
 DE
 XX
 XX
 KM Nodavirus capsid protein; chimeric protein; anti-parallel beta barrel;
 KM immune response; chimeric virus-like particle; gene-delivery vector;
 KM hepatitis B infection; vaccine; vesicular stomatitis viral infection;
 KM respiratory syncytial virus; malaria; Flock House virus; FHV.
 XX
 XX Synthetic.
 OS
 XX MO9929723-A1.
 PM
 XX 17-JUN-1999.
 PD
 XX 07-DEC-1998; 98WO-US25922.
 PF
 XX 08-DEC-1997; 97US-0986659.
 PR
 XX (PENT-) PENTAMER PHARM.
 PA (SCRI) SCRIPPS RES INST.
 XX
 PI Hall SG;
 DR WPI; 1999-385574/32.
 XX
 PT Recombinant chimeric nodavirus particles
 XX
 PS Disclosure; Page 62-63; 69pp; English.
 XX
 CC This sequence represents a peptide used in the construction of a
 CC flock house virus (FHV) chimeric particle.
 CC The invention relates to a chimeric protein comprising a nodavirus capsid
 CC protein free from deletions, having a core structure constituted by
 CC anti-parallel beta barrels, and a heterologous peptide segment situated
 CC between a pair of strands of one of the beta barrels. The chimeric
 CC protein is used to induce an immune response in an animal. The chimeric
 CC proteins can be assembled to form chimeric virus-like particles that are
 CC useful in therapeutic applications, such as vaccines and gene-delivery
 CC vectors, and in diagnostic applications, such as kits for the testing of
 CC body tissue or fluid samples. The chimeric virus-like particles mimic
 CC infectious viruses and parasites and are useful for treating hepatitis B
 CC infection, vesicular stomatitis viral infection, bovine and human
 CC respiratory syncytial virus, as well as malaria. Flock House virus (FHV)
 CC is a non-pathogenic nodavirus that can be used to genetically engineer
 CC virus-like particles carrying antigenic peptides on their surface. The
 CC FHV capsid protein has a remarkable functional versatility. A region of
 CC the capsid protein is amenable to insertion of heterologous peptide
 CC segments without affecting assembly of the viral coat or capsid.
 CC
 XX
 SQ Sequence 23 AA;
 Query Match 100.0%; Score 103; DB 20; Length 23;
 Best Local Similarity 100.0%; Pred. No. 1.2e-07;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EMSPCSVTCGNGIOVRIK 18
 DB 1 EMSPCSVTCGNGIOVRIK 18
 RESULT 5
 AAY22139
 ID AAY22139 standard; peptide; 23 AA.
 XX
 XX AAY22139;
 AC
 XX 08-SEP-1999 (first entry)
 DT
 XX P. falciparum CSP:IX liver specific ligand.
 DE
 XX Nodavirus capsid protein; chimeric protein; anti-parallel beta barrel;
 KM

KM immune response; chimeric virus-like particle; gene-delivery vector;
 KM hepatitis B infection; vaccine; vesicular stomatitis viral infection;
 KM respiratory syncytial virus; malaria; Flock House virus; FHV.
 XX
 XX Plasmodium falciparum.
 OS
 XX MO9929723-A1.
 PM
 XX 17-JUN-1999.
 PD
 XX 07-DEC-1998; 98WO-US25922.
 PF
 XX 08-DEC-1997; 97US-0986659.
 PR
 XX (PENT-) PENTAMER PHARM.
 PA (SCRI) SCRIPPS RES INST.
 XX
 PI Hall SG;
 DR WPI; 1999-385574/32.
 XX
 PT Recombinant chimeric nodavirus particles
 XX
 PS Claim 3; Page 57; 69pp; English.
 XX
 CC This sequence represents a Plasmodium falciparum CSP:IX liver
 CC specific ligand.
 CC The invention relates to a chimeric protein comprising a nodavirus capsid
 CC protein free from deletions, having a core structure constituted by
 CC anti-parallel beta barrels, and a heterologous peptide segment situated
 CC between a pair of strands of one of the beta barrels. The chimeric
 CC protein is used to induce an immune response in an animal. The chimeric
 CC proteins can be assembled to form chimeric virus-like particles that are
 CC useful in therapeutic applications, such as vaccines and gene-delivery
 CC vectors, and in diagnostic applications, such as kits for the testing of
 CC body tissue or fluid samples. The chimeric virus-like particles mimic
 CC infectious viruses and parasites and are useful for treating hepatitis B
 CC infection, vesicular stomatitis viral infection, bovine and human
 CC respiratory syncytial virus, as well as malaria. Flock House virus (FHV)
 CC is a non-pathogenic nodavirus that can be used to genetically engineer
 CC virus-like particles carrying antigenic peptides on their surface. The
 CC FHV capsid protein has a remarkable functional versatility. A region of
 CC the capsid protein is amenable to insertion of heterologous peptide
 CC segments without affecting assembly of the viral coat or capsid.
 CC
 XX
 SQ Sequence 23 AA;
 Query Match 100.0%; Score 103; DB 20; Length 23;
 Best Local Similarity 100.0%; Pred. No. 1.2e-07;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EMSPCSVTCGNGIOVRIK 18
 DB 1 EMSPCSVTCGNGIOVRIK 18
 RESULT 6
 AAY22154
 ID AAY22154 standard; Protein; 33 AA.
 XX
 XX AAY22154;
 AC
 XX 08-SEP-1999 (first entry)
 DT
 XX Peptide used for FHV chimeric particle construction.
 DE
 XX Nodavirus capsid protein; chimeric protein; anti-parallel beta barrel;
 KM immune response; chimeric virus-like particle; gene-delivery vector;
 KM hepatitis B infection; vaccine; vesicular stomatitis viral infection;
 KM respiratory syncytial virus; malaria; Flock House virus; FHV.
 XX
 XX Synthetic.
 OS

PN WO9929723-A1.
 XX
 PD 17-JUN-1999.
 XX
 PF 07-DEC-1998; 98WO-US35922.
 XX
 PR 08-DEC-1997; 97US-096659.
 XX
 PA (PENT-) PENTAMER PHARM.
 PA (SCRI) SCRIPPS RES INST.
 XX
 PI Hall SG;
 XX
 XX WPI; 1999-385574/32.
 XX
 XX
 PT Recombinant chimeric nodavirus particles
 XX
 PS Disclosure; Page 63; 69pp; English.
 XX
 CC This sequence represents a peptide used in the construction of a
 CC flock house virus (FHV) chimeric particle.
 CC The invention relates to a chimeric protein comprising a nodavirus capsid
 CC protein free from deletions, having a core structure constituted by
 CC anti-parallel beta barrels, and a heterologous peptide segment situated
 CC between a pair of strands of one of the beta barrels. The chimeric
 CC protein is used to induce an immune response in an animal. The chimeric
 CC proteins can be assembled to form chimeric virus-like particles that are
 CC useful in therapeutic applications, such as vaccines and gene-delivery
 CC vectors, and in diagnostic applications, such as kits for the testing of
 CC body tissue or fluid samples. The chimeric virus-like particles mimic
 CC infectious viruses and parasites and are useful for treating hepatitis B
 CC infection, vesicular stomatitis viral infection, bovine and human
 CC respiratory syncytial virus, as well as malaria. Flock House virus (FHV)
 CC is a non-pathogenic nodavirus that can be used to genetically engineer
 CC virus-like particles carrying antigenic peptides on their surface. The
 CC FHV capsid protein has a remarkable functional versatility. A region of
 CC the capsid protein is amenable to insertion of heterologous peptide
 CC segments without affecting assembly of the viral coat or capsid.
 XX
 Q0 Sequence 33 AA;

	Query March -	Score 103;	DB 20;	Length 33;
	Best Local Similarity	100.0%;	Pred. No. 1.7e-07;	
Matches	18; Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
QY	1 EWSPCSVTCGANGICQYRIK 18 			
Db	6 EWSPCSVTCTGGNGIQYRIK 23			
RESULT 7				
ID	AAMS9274			
XX	AAMS9274 standard; peptide; 46 AA.			
AC	AAMS9274;			
XX				
DT	27-AUG-1998 (first entry)			
DE	P. falciparum circumsporozoite region II peptide fragment CS271C-His6.			
KX				
KW	Malaria; gene therapy; hepatocyte; liver; circumsporozoite; ligand; targeted delivery; therapy; disease; cancer; hepatitis; cystic fibrosis; hypercholesterolaemia; phenylketonuria; haemophilia.			
OS	Plasmodium falciparum.			
XX				
FN	US5766899-A.			
XX				
PD	16-JUN-1998.			
XX				
PF	27-FEB-1995; 95US-0395602.			
XX				
PR	27-FEB-1995; 95US-0395602.			

```

PA (TEXTA ) UNIV TEXAS SYSTEM.
XX
XX Ding Z, Kuo MT;
XX
XX WPI; 1998-361692/31.
XX
XX Complexes for targetted delivery of nucleic acids to hepatocytes -
XX containing Plasmodium circumsporozoite polypeptide as targeting
XX ligand
XX
XX Claim 2; Column 31-32; 34pp; English.
XX
XX
XX AAM59270-W59374 are fragments of malarial circumsporozoite (CS) region
XX It isolated from Plasmodium falciparum. These fragments can be used as
XX ligands in a method for the targetted delivery of nucleic acid to cells
XX in culture or cells in vivo, especially where the cells are hepatocytes.
XX Therapy of diseases such as cancer, malaria, hepatitis, cystic fibrosis,
XX hypercholesterolaemia, phenylketonuria and haemophilia is mentioned.
XX CS polypeptides are liver cell specific with rapid hepatic invasion.
XX They are more efficient than the prior art asialoorosomucoid (ASOR)
XX ligands, of which there may be an accumulation in certain diseases due to
XX receptor downregulation.
XX
XX Sequence 46 AA:
XX
XX Query Match 100.0%; Score 103; DB 19; Length 46;
XX Best Local Similarity 100.0%; Pred. No. 2,4e-07;
XX Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0.
XX
XX 1 EWSPCSVTCGNGIOVRK 18
XX |||||
XX 29 EWSPCSVTCGNGIOVRK 46
XX
XX RESULT 8
XX AAR71644
XX ID AAR71644 standard; Peptide; 54 AA.
XX
XX AC AAR71644;
XX
XX DT 25-MAR-2003 (updated)
XX DT 19-SEP-1995 (first entry)
XX
XX DE Circumsporozoite Region II+.
XX
XX CM Circumsporozoite; Region II+; malaria; hepatocyte; receptor.
XX
XX OS Plasmodium falciparum.
XX
XX FH Key Location/Qualifiers
XX FT Peptide 20..37
XX FT /label= Region-II+
XX
XX FN WO9507094-A1.
XX
XX PD 16-MAR-1995.
XX
XX PF 09-SEP-1994; 94WO-US10186.
XX
XX PR 10-SEP-1993; 93US-0119694.
XX
XX PA (UUNY ) UNIV NEW YORK STATE.
XX
XX PI Cerami C, Frevert U, Nussenzweig V, Simms P;
XX
XX WPI; 1995-123236/16.
XX
XX New peptide inhibitors of circumsporozoite binding - used to
XX develop prods. for preventing malaria infection or for delivering
XX substances to hepatocyte(s)
XX
XX Disclosure; Page 14; 13pp; English.

```

```

XX XX The sequence surrounding Region II+ of P. falciparum
CC CC circumsporozoite protein (CSP) is given in AAR71644. Region II+,
CC encompassing a hepatocyte ligand for malaria sporozoites is
CC located at the C-terminus of CSP. Mimetic peptides (AAR71646-54)
CC based on Region II+ were tested for their ability to prevent or
CC attenuate malaria infection.
CC (Updated on 25-MAR-2003 to correct PN field.)
CC
SQ Sequence 54 AA:
Query Match 100.0%; Score 103; DB 16; Length 54;
Best Local Similarity 100.0%; Pred. No. 2.9e-07;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EMSPCSVTCGNGIOVRK 18
DB 20 EMSPCSVTCGNGIOVRK 37

RESULT 9
AAM59273
ID AAM59273 standard; peptide; 126 AA.
XX
XX AAM59273;
AC
XX
XX 27-AUG-1998 (first entry)
DT
XX
XX P. falciparum circumsporozoite region II peptide fragment CS2771VC-His6.
DE
XX
XX Malaria; gene therapy; hepatocyte; liver; circumsporozoite; ligand;
KW targeted delivery; therapy; disease; cancer; hepatitis; cystic fibrosis;
KW hypercholesterolemia; phenylketonuria; haemophilia.
XX
XX Plasmodium falciparum.
OS
XX
XX US5766899-A.
PN
XX
XX 16-JUN-1998.
PD
XX
XX 27-FEB-1995; 95US-0395602.
PF
XX
XX 27-FEB-1995; 95US-0395602.
PR
XX
XX 27-FEB-1995; 95US-0395602.
PA
XX (TEXA ) UNIV TEXAS SYSTEM.
PI
XX Ding Z, Kuo MT;
DR WPI; 1998-361692/31.
XX
XX Complexes for targeted delivery of nucleic acids to hepatocytes -
PT containing Plasmodium circumsporozoite polypeptide as targeting
PT ligand
XX
XX Claim 2; Column 31-32; 34pp; English.
PS
XX AAM59270-M59274 are fragments of malarial circumsporozoite (CS) region
CC II isolated from Plasmodium falciparum. These fragments can be used as
CC ligands in a method for the targeted delivery of nucleic acid to cells
CC in culture or cells in vivo, especially where the cells are hepatocytes.
CC Therapy of diseases such as cancer, malaria, hepatitis, cystic fibrosis,
CC hypercholesterolemia, phenylketonuria and haemophilia is mentioned.
CC CS polypeptides are liver cell specific with rapid hepatic invasion.
CC They are more efficient than the prior art asialosomucoid (ASOR).
CC ligands, of which there may be an accumulation in certain diseases due to
CC receptor downregulation.
CC
SQ Sequence 126 AA:
Query Match 100.0%; Score 103; DB 19; Length 126;
Best Local Similarity 100.0%; Pred. No. 6.6e-07;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 EMSPCSVTCGNGIOVRK 18
DB 109 EMSPCSVTCGNGIOVRK 126

RESULT 10
AAR07290
ID AAR07290 standard; protein; 180 AA.
XX
XX AAR07290;
AC
XX
XX 25-MAR-2003 (updated)
DT
XX 28-JAN-1991 (first entry)
DE
XX Circumsporozoite analogue Falciparum 4.
XX
XX CS protein; plasmodium; malaria; vaccine.
KW
XX Synthetic.
OS
XX
XX EP392820-A.
PN
XX
XX 17-OCT-1990.
PD
XX
XX 11-APR-1990; 90EP-0303907.
PF
XX
XX 11-APR-1989; 89US-0336288.
PR
XX
XX (CHIR ) CHIRON CORP.
PA
XX
XX Barr PY, Bathurst IC, Gibson HU,
PI
XX WPI; 1990-314486/42.
DR
XX N-PSDB; AA006153.
XX
XX Recombinant plasmodium circumsporozoite analogues - lacking
PT one or more repeat epitope(s) for use as a malaria sub-unit
PT vaccine.
XX
XX Claim 10; Fig 10; 22pp; English.
PS
XX
XX The protein was produced by expression of a synthetic gene. The
CC analogue comprises AAs 68-123 of the native P. falciparum CS
CC protein, followed by four repeat sequences (three "B"rs, i.e. NNMP)
CC and one "A", i.e. NNDP), followed by AAs 289-392 of the native
CC protein. Reduction of the immunological dominance of the repeats
CC relative to the epitopes in the regions flanking the repeats
CC enhances sporozoite neutralising activity.
CC See also AAR07287-91.
CC
XX (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ Sequence 180 AA:
Query Match 100.0%; Score 103; DB 11; Length 180;
Best Local Similarity 100.0%; Pred. No. 9.5e-07;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EMSPCSVTCGNGIOVRK 18
DB 134 EMSPCSVTCGNGIOVRK 151

RESULT 11
AAR07289
ID AAR07289 standard; protein; 184 AA.
XX
XX AAR07289;
AC
XX
XX 25-MAR-2003 (updated)
DT
XX 28-JAN-1991 (first entry)
DE
XX Circumsporozoite analogue Falciparum 3.
XX

```

KM	CS protein; plasmodium; malaria; vaccine.
XX	
OS	Synthetic.
XX	
PN	EP392820-A.
XX	
PD	17-OCT-1990.
XX	
PF	11-APR-1990; 90EP-0303907.
XX	
PR	11-APR-1989; 89US-0336288.
XX	
PA	(CHIR) CHIRON CORP.
XX	
PI	Barr PU, Bathurst IC, Gibson HL;
XX	
DR	WPI; 1990-314486/42.
DR	N-PSDB; AAQ06152.
XX	
PT	Recombinant plasmodium circumsporozoite analogues - lacking
PT	one or more repeat epitope(s) for use as a malaria sub-unit
PT	vaccine.
XX	
PS	Claim 10; Fig 8; 22pp; English.
XX	
CC	The protein was produced by expression of a synthetic gene. The
CC	analogue comprises AAs 43-123 of the native P. falciparum CS
CC	protein, followed by four repeat sequences (three "B"s, i.e. NANP)
CC	and one "A", i.e. NNDP), followed by AAs 289-374 of the native
CC	protein. Reduction of the immunological dominance of the repeats
CC	relative to the epitopes in the regions flanking the repeats
CC	enhances sporozoite neutralising activity.
CC	(See also AAR07287-91.
CC	Updated on 25-MAR-2003 to correct PA field.)
XX	
SQ	Sequence 184 AA:
QY	1 EMSPCVCNGNGIOVRIK 18
DB	156 EMSPCVCNGNGIOVRIK 173
RESULT 12	
AAR13175	
ID	AAR13175 standard; Protein; 309 AA.
XX	
AC	AAR13175;
XX	
DT	25-MAR-2003 (updated)
DT	29-AUG-1991 (first entry)
XX	
DE	NS1_81-RLfdelta9.
XX	
KM	Immunogenic determinant; circumsporozoite; CS; vaccine; malaria;
KM	hybrid; influenza virus; non-structural protein 1; fusion.
XX	
OS	Plasmodium falciparum.
OS	Influenza virus (A/PR/8/34/).
XX	
PH	Key
FT	Location/Qualifiers
FT	1..81
FT	/label= N-terminal of NS1
FT	/note= "Influenza virus nonstructural protein 1"
FT	82..87
FT	/label= synthetic linker
FT	88
FT	/label= artifact
FT	/note= "see comments"
FT	89..193
FT	Region

FT	/label= Aas 19-123 of CS protein
PT	/note= "Region 1 contg. flanking region less
FT	signal sequence"
FT	194..309
FT	/label= Aas 297-412 of CS protein
FT	/note= "Region II flanking region minus 9 N-term-
XX	inal Aas"
PV	EP432965-A.
XX	19-JUN-1991.
PD	
XX	06-DEC-1990; 90EP-0313257.
PF	
XX	08-DEC-1989; 89US-0447746.
PR	
PA	(SMIK) SMITHKLINE BEECHAM.
PA	(USSA) US SEC OF ARMY.
PA	(BIOM-) BIOMEDICAL RES INST.
PA	(GROS/) GROSS M S.
XI	
P1	Gross MS, Gordon DM, Hollingdale MR;
XX	
DR	WPI; 1991-179771/25.
XX	
PT	- For vaccine against malaria infection in humans.
XX	
PS	Example 1; Page 7; 18pp; English.
XX	
CC	The polypeptide is prepd. by genetic engineering of genes encoding
CC	the P. falciparum circumsporozoite (CS) protein [Dane et al.,
CC	Science 225 : 553 (1984)], and the influenza virus non-structural
CC	protein 1 (NS1), [Baez et al., Nucleic Acids Research, 8 : 5845
CC	(1980)]. The DNA encoding the 1st 81 AAs of the N-terminal of NS1
CC	(NS1 81) is linked via a synthetic sequence to DNA encoding Region
CC	I contg. flanking regionless the 18 AA signal region, which in
CC	turn is fused to DNA encoding Region II contg. flanking region
CC	less the first nine N-terminal AAs. This CS fusion is designated
CC	Rtldelta9. The Pro residue separating the Asp (at the C-terminal
CC	of the linker) from Rtldelta9 is an artifact of a filled in BamHI
CC	site. The peptide can be used in a vaccine for protection against
CC	malaria.
CC	See also AAR12306-R12311 and AAR13176-R13179.
CC	(Updated on 25-MAR-2003 to correct PA field.)
XX	
SQ	Sequence 309 AA;
	Query Match 100.0%; Score 103; DB 12; Length 309;
	Best Local Similarity 100.0%; Prd. NO. 1.6e-06;
	Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0,
OY	1 EMSPCSVTCGGNGIOVR1K 18
DB	242 EMSPCSVTCGGNGIOVR1K 259
RESULT 13	
AAR07945	
ID	AAR07945 standard; protein, 319 AA.
XX	
AC	AAR07945;
XX	
DT	25-MAR-2003 (updated)
DT	22-FEB-1991 (first entry)
XX	
DE	NS181RIFAuth plaemid product.
XX	
KM	Malaria; vaccine.
XX	
OS	Plasmodium falciparum.
XX	
FH	Key Location/Qualifiers


```

FT Domain 1..81
FT /label= NS181 protein fragment
FT /note= "from plasmid pMG-1"
FT 89..193
FT Domain /label= Fragment of circumsporozoite protein
FT 204..319
FT /label= Fragment of circumsporozoite protein
XX EP398540-A.
XX 22-NOV-1990.
XX 01-MAY-1990; 90EP-0304720.
XX 03-MAY-1989; 89US-0346863.
XX (SMIK ) SMITHKLINE BEECHAM.
XX (GROSS/) GROSS M S.
XX Gross MS, Young JF;
XX WPI, 1990-350299/47.
XX N-PSDB; AA006580.
XX New polypeptide used in malaria vaccine - comprises immunogenic
XX determinants from 1st and 2nd flanking regions of Plasmodium
XX surface protein and intermediate repeat domain
XX Example 2; Page 11-12; 24pp; English.
XX The product is useful in preparation of vaccines for treatment and
XX prophylaxis of plasmodium sporozite infection. It may be easily
XX produced in large pure quantities from a transformed E.coli
XX expression system.
XX (Updated on 25-MAR-2003 to correct PA field.)
XX Sequence 319 AA;
SQ
Query Match 100.0%; Score 103; DB 11; Length 319;
Best Local Similarity 100.0%; Pred. No. 1.7e-06;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 EWSPCSVTCGNGIQVRIK 18
DB 252 EWSPCSVTCGNGIQVRIK 269
RESULT 14
AAR13176 AAR13176 standard; Protein; 319 AA.
XX AAR13176;
XX 25-MAR-2003 (updated)
DT 29-AUG-1991 (first entry)
XX NS1_81-RLFAuth.
XX Immunogenic determinant; circumsporozoite; CS; vaccine; malaria;
XX hybrid; Influenza virus; non-structural protein 1; fusion.
XX Plasmodium falciparum.
XX Influenza virus (A/PR/8/34).
XX Key Location/Qualifiers
XX 1..81
XX /label= N-terminal of NS1
XX /note= "Influenza virus nonstructural protein 1"
XX Peptide 82..87
XX /label= synthetic linker
XX 88
XX /label= artifact
XX /note= "see comments"

```

```

FT Region 89..193
FT /label= Aas 19-123 of CS protein
FT /note= "Region I contg. flanking region less
FT signal sequence"
FT 194
FT /label= artifact
FT /note= "see comments"
FT 195..319
FT /label= Aas 288-412 of CS protein
FT /note= "Region II flanking region"
XX EP432965-A.
XX 19-JUN-1991.
XX 06-DEC-1990; 90EP-0313257.
XX 08-DEC-1989; 89US-0447746.
XX (SMIK ) SMITHKLINE BEECHAM.
XX (USSA ) US SEC OF ARMY.
XX (BIOM-) BIOMEDICAL RES INST.
XX (GROSS/) GROSS M S.
XX Gross MS, Gordon DM, Hollingdale MR;
XX WPI, 1991-179771/25.
XX Polypeptide comprising immunogenic determinants from P falciparum
XX - for vaccine against malaria infection in humans.
XX Example 2; Page 10; 18pp; English.
XX The polypeptide is prepd. by genetic engineering of genes encoding
XX the P. falciparum circumsporozoite (CS) protein [Dame et al.,
XX Science 225 : 593 (1984)], and the Influenza virus non-structural
XX protein 1 (NS1) [Baez et al., Nucleic Acids Research, 8 : 5845
XX (1980)]. The DNA encoding the 1st 81 AAs of the N-terminal of NS1
XX (NS1_81) is linked via a synthetic sequence to DNA encoding Region
XX I contg. flanking regionless the 18 AA signal region, which in
XX turn is fused to DNA encoding Region II-contg. flanking region.
XX This CS fusion is designated RLFAuth. The Pro residue separating
XX the Asp (at the C-terminal of the linker) from RLFAuth is an arti-
XX fact of a filled in BamHI site; the Gly separating Region I and
XX Region II-contg. CS flanking regions is an artifact of a synthetic
XX FokI/NotI linker. The peptide can be used in a vaccine for
XX protection against malaria.
XX The complete nucleotide and AA sequences are given in EP-304720,
XX CC filed May 1, 1990.
XX See also AAR12306-R12311 and AAR13175-R13179.
XX (Updated on 25-MAR-2003 to correct PA field.)
XX Sequence 319 AA;
SQ
Query Match 100.0%; Score 103; DB 12; Length 319;
Best Local Similarity 100.0%; Pred. No. 1.7e-06;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 EWSPCSVTCGNGIQVRIK 18
DB 252 EWSPCSVTCGNGIQVRIK 269
RESULT 15
AAR13177 AAR13177 standard; Protein; 327 AA.
XX AAR13177;
XX 25-MAR-2003 (updated)
DT 29-AUG-1991 (first entry)
XX NS1_81-RLFAuth + (NANP)2.

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 23, 2003, 15:33:08 ; Search time 21.75 Seconds

(without alignments)
35.016 Million cell updates/sec

Title: US-09-980-564-13

Perfect score: 103

Sequence: 1 EWSPCSTCGNGIQVRIK 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Issued Patents, AA:*

2: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*

3: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*

4: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*

5: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*

6: /cgn2_6/ptodata/1/1aa/6C_COMB.pep:*

7: /cgn2_6/ptodata/1/1aa/6D_COMB.pep:*

8: /cgn2_6/ptodata/1/1aa/6E_COMB.pep:*

9: /cgn2_6/ptodata/1/1aa/6F_COMB.pep:*

10: /cgn2_6/ptodata/1/1aa/6G_COMB.pep:*

11: /cgn2_6/ptodata/1/1aa/6H_COMB.pep:*

12: /cgn2_6/ptodata/1/1aa/6I_COMB.pep:*

13: /cgn2_6/ptodata/1/1aa/6J_COMB.pep:*

14: /cgn2_6/ptodata/1/1aa/6K_COMB.pep:*

15: /cgn2_6/ptodata/1/1aa/6L_COMB.pep:*

16: /cgn2_6/ptodata/1/1aa/6M_COMB.pep:*

17: /cgn2_6/ptodata/1/1aa/6N_COMB.pep:*

18: /cgn2_6/ptodata/1/1aa/6O_COMB.pep:*

19: /cgn2_6/ptodata/1/1aa/6P_COMB.pep:*

20: /cgn2_6/ptodata/1/1aa/6Q_COMB.pep:*

21: /cgn2_6/ptodata/1/1aa/6R_COMB.pep:*

22: /cgn2_6/ptodata/1/1aa/6S_COMB.pep:*

23: /cgn2_6/ptodata/1/1aa/6T_COMB.pep:*

24: /cgn2_6/ptodata/1/1aa/6U_COMB.pep:*

25: /cgn2_6/ptodata/1/1aa/6V_COMB.pep:*

26: /cgn2_6/ptodata/1/1aa/6W_COMB.pep:*

27: /cgn2_6/ptodata/1/1aa/6X_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	103	100.0	18	US-08-395-602A-1	Sequence 1, Appl
2	103	100.0	18	US-08-021-625D-1	Sequence 1, Appl
3	103	100.0	23	US-08-395-602A-2	Sequence 2, Appl
4	103	100.0	23	US-08-021-625D-2	Sequence 2, Appl
5	103	100.0	23	US-08-986-659B-9	Sequence 9, Appl
6	103	100.0	23	US-08-986-659B-33	Sequence 33, Appl
7	103	100.0	33	US-08-986-659B-34	Sequence 34, Appl
8	103	100.0	46	US-08-395-602A-5	Sequence 5, Appl
9	103	100.0	46	US-08-021-625D-5	Sequence 5, Appl
10	103	100.0	126	US-08-395-602A-4	Sequence 4, Appl
11	103	100.0	126	US-08-021-625D-4	Sequence 4, Appl
12	103	100.0	412	US-08-313-288B-18	Sequence 18, Appl
13	103	100.0	423	US-08-760-797A-1	Sequence 1, Appl
14	103	100.0	424	US-08-760-797A-3	Sequence 3, Appl
15	103	100.0	424	US-08-932-929B-1	Sequence 1, Appl
16	103	100.0	424	US-08-932-929B-3	Sequence 3, Appl
17	83	80.6	20	US-08-986-659B-8	Sequence 8, Appl
18	77	74.8	378	5171843-9	Patent No. 5171843
19	77	74.8	525	US-09-369-364A-21	Sequence 21, Appl
20	76	73.8	478	US-08-155-888-2	Sequence 2, Appl
21	73	70.9	114	US-08-309-604-2	Sequence 2, Appl
22	73	70.9	559	US-08-313-288B-14	Sequence 14, Appl
23	71	68.9	345	US-09-182-145-11	Sequence 11, Appl
24	71	68.9	367	US-09-182-145-12	Sequence 12, Appl
25	70	68.0	568	US-07-862-021B-14	Sequence 14, Appl
26	70	68.0	568	PCT-US93-03164-14	Sequence 14, Appl
27	70	68.0	1205	US-09-491-522-11	Sequence 11, Appl

28	70	68.0	1211	4	US-09-491-522-5	Sequence 5, Appl
29	69	67.0	23	1	US-07-646-531D-5	Sequence 5, Appl
30	69	67.0	23	1	US-07-646-531D-6	Sequence 6, Appl
31	69	67.0	23	2	US-08-488-273-5	Sequence 5, Appl
32	69	67.0	23	2	US-08-488-273-6	Sequence 6, Appl
33	69	67.0	23	4	US-09-197-770B-11	Sequence 11, Appl
34	69	67.0	23	4	US-09-197-770B-12	Sequence 12, Appl
35	69	67.0	23	6	5426100-5	Patent No. 5426100
36	69	67.0	23	6	5426100-6	Patent No. 5426100
37	69	67.0	51	2	US-08-799-173A-11	Sequence 11, Appl
38	69	67.0	53	1	US-07-862-021B-18	Sequence 18, Appl
39	69	67.0	53	1	PCT-US93-03164-18	Sequence 18, Appl
40	69	67.0	60	1	US-07-646-531D-12	Sequence 12, Appl
41	69	67.0	60	2	US-08-488-273-12	Sequence 12, Appl
42	69	67.0	60	6	5426100-12	Patent No. 5426100
43	69	67.0	218	3	US-08-985-526-1	Sequence 1, Appl
44	69	67.0	239	5	PCT-US93-01652-1	Sequence 1, Appl
45	69	67.0	441	3	US-08-985-526-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1
US-08-395-602A-1

; Sequence 1, Application US/08395602A

; Patent No. 5766899

; GENERAL INFORMATION:

; APPLICANT: Kuo, M. Tien

; TITLE OF INVENTION: Targeted Nucleic Acid Delivery Into

; NUMBER OF SEQUENCES: 5

; CORRESPONDENCE ADDRESS:

; ADDRESS: Arnold, White & Durkee

; STREET: P.O. Box 4433

; CITY: Houston

; STATE: Texas

; COUNTRY: United States of America

; ZIP: 77210

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent in Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/395,602A

; FILING DATE:

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Highlander, Steven L.

; REGISTRATION NUMBER: 37,642

; TELEPHONE: (512) 418-3000

; TELEFAX: (713) 789-2679

; TRIEX: 79-0924

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 18 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; US-08-395-602A-1

Query Match 100.0%; Score 103; DB 1; Length 18;

Best Local Similarity 100.0%; Pred. No. 9.5e-09;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EWSPCSTCGNGIQVRIK 18

Db 1 EWSPCSTCGNGIQVRIK 18

RESULT 2
US-08-021-625D-1
Sequence 1, Application US/08021625D
Patent No. 5976851
GENERAL INFORMATION:
APPLICANT: Kuo, M. Tien
APPLICANT: Ding, Zhi-Ming
TITLE OF INVENTION: Targeted Nucleic Acid Delivery into
TITLE OF INVENTION: Liver Cells
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: United States of America
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/021,625D
FILING DATE: 16-FEB-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Highlander, Steven L.
REGISTRATION NUMBER: 37,642
REFERENCE/DOCKET NUMBER: UTSC:410/HYL
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
TELEX: 79-0924
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 18 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-021-625D-1
Query Match 100.0%; Score 103; DB 2; Length 18;
Best Local Similarity 100.0%; Pred. No. 9.5e-09;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EMSPCSVTCGNGIQVRIK 18
|||||
Db 1 EMSPCSVTCGNGIQVRIK 18

RESULT 3
US-08-395-602A-2
Sequence 2, Application US/08395602A
Patent No. 5766899
GENERAL INFORMATION:
APPLICANT: Kuo, M. Tien
APPLICANT: Ding, Zhi-Ming
TITLE OF INVENTION: Targeted Nucleic Acid Delivery into
TITLE OF INVENTION: Liver Cells
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: United States of America
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/395,602A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Highlander, Steven L.
REGISTRATION NUMBER: 37,642
REFERENCE/DOCKET NUMBER: UTSC:410/HYL
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
TELEX: 79-0924
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 23 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-395-602A-2
Query Match 100.0%; Score 103; DB 1; Length 23;
Best Local Similarity 100.0%; Pred. No. 1.2e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EMSPCSVTCGNGIQVRIK 18
|||||
Db 1 EMSPCSVTCGNGIQVRIK 18

RESULT 4
US-08-021-625D-2
Sequence 2, Application US/08021625D
Patent No. 5976851
GENERAL INFORMATION:
APPLICANT: Kuo, M. Tien
APPLICANT: Ding, Zhi-Ming
TITLE OF INVENTION: Targeted Nucleic Acid Delivery into
TITLE OF INVENTION: Liver Cells
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: United States of America
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/021,625D
FILING DATE: 16-FEB-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Highlander, Steven L.
REGISTRATION NUMBER: 37,642
REFERENCE/DOCKET NUMBER: UTSC:410/HYL
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
TELEX: 79-0924
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 23 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-021-625D-2
Query Match 100.0%; Score 103; DB 2; Length 23;

Best Local Similarity 100.0%; Pred. No. 1.2e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EWSPCSTCGNGIOVRK 18
Db 1 EWSPCSTCGNGIOVRK 18

RESULT 5
US-08-986-659B-9

Sequence 9, Application US/08986659B

Patent No. 6171591

GENERAL INFORMATION:

APPLICANT: Hall, Stephen G.

TITLE OF INVENTION: RECOMBINANT NODAVIRUS RELATED

NUMBER OF SEQUENCES: 46

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Olson & Hierl, Ltd.

STREET: 20 No. 6171591th Wacker Drive, 36th Floor

CITY: Chicago

STATE: IL

COUNTRY: USA

ZIP: 60606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/986,659B

FILING DATE: 08-DEC-1997

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Talivaldis Cepuritis

REGISTRATION NUMBER: 20,818

REFERENCE/DOCKET NUMBER: 549.0

TELECOMMUNICATION INFORMATION:

TELEPHONE: 312-580-1180

TELEFAX: 312-580-1189

INFORMATION FOR SEQ ID NO: 9:

SEQUENCE CHARACTERISTICS:

LENGTH: 23 amino acids

TYPE: amino acid

STRANDEDNESS: N/A

TOPOLOGY: unknown

MOLECULE TYPE: peptide

US-08-986-659B-9

OY 1 EWSPCSTCGNGIOVRK 18
Db 1 EWSPCSTCGNGIOVRK 18

RESULT 6

US-08-986-659B-33

Sequence 33, Application US/08986659B

Patent No. 6171591

GENERAL INFORMATION:

APPLICANT: Hall, Stephen G.

TITLE OF INVENTION: RECOMBINANT NODAVIRUS RELATED

NUMBER OF SEQUENCES: 46

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Olson & Hierl, Ltd.

STREET: 20 No. 6171591th Wacker Drive, 36th Floor

CITY: Chicago

STATE: IL

COUNTRY: USA

ZIP: 60606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/986,659B

FILING DATE: 08-DEC-1997

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Talivaldis Cepuritis

REGISTRATION NUMBER: 20,818

REFERENCE/DOCKET NUMBER: 549.0

TELECOMMUNICATION INFORMATION:

TELEPHONE: 312-580-1180

TELEFAX: 312-580-1189

INFORMATION FOR SEQ ID NO: 33:

SEQUENCE CHARACTERISTICS:

LENGTH: 23 amino acids

TYPE: amino acid

STRANDEDNESS: N/A

TOPOLOGY: unknown

MOLECULE TYPE: peptide

US-08-986-659B-33

OY 1 EWSPCSTCGNGIOVRK 18
Db 1 EWSPCSTCGNGIOVRK 18

RESULT 7
US-08-986-659B-34

Sequence 34, Application US/08986659B

Patent No. 6171591

GENERAL INFORMATION:

APPLICANT: Hall, Stephen G.

TITLE OF INVENTION: RECOMBINANT NODAVIRUS RELATED

NUMBER OF SEQUENCES: 46

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Olson & Hierl, Ltd.

STREET: 20 No. 6171591th Wacker Drive, 36th Floor

CITY: Chicago

STATE: IL

COUNTRY: USA

ZIP: 60606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/986,659B

FILING DATE: 08-DEC-1997

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Talivaldis Cepuritis

REGISTRATION NUMBER: 20,818

REFERENCE/DOCKET NUMBER: 549.0

TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-580-1180
TELEFAX: 312-580-1189
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 33 amino acids
TYPE: amino acid
STRANDEDNESS: N/A
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-986-659B-34

Query Match 100.0%; Score 103; DB 3; Length 33;
Best Local Similarity 100.0%; Pred. No. 1.8e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EMSPCSVTCGNGIOVRK 18
Db 6 EMSPCSVTCGNGIOVRK 23

RESULT 8

US-08-395-602A-5
Sequence 5, Application US/08395602A
Patent No. 5766899
GENERAL INFORMATION:
APPLICANT: Kuo, M. Tien
APPLICANT: Ding, Zhi-ming
TITLE OF INVENTION: Targeted Nucleic Acid Delivery into
TITLE OF INVENTION: Liver Cells
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: United States of America
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/395,602A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Highlander, Steven L.
REGISTRATION NUMBER: 37,642
REFERENCE/DOCKET NUMBER: UTSC:410/HYL
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
TELEX: 79-0924
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 46 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-395-602A-5

Query Match 100.0%; Score 103; DB 1; Length 46;
Best Local Similarity 100.0%; Pred. No. 2.5e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EMSPCSVTCGNGIOVRK 18
Db 29 EMSPCSVTCGNGIOVRK 46

RESULT 9

US-08-021-625D-5
Sequence 5, Application US/08021625D
Patent No. 5976851
GENERAL INFORMATION:
APPLICANT: Kuo, M. Tien
APPLICANT: Ding, Zhi-ming
TITLE OF INVENTION: Targeted Nucleic Acid Delivery into
TITLE OF INVENTION: Liver Cells
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: United States of America
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/021,625D
FILING DATE: 16-FEB-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Highlander, Steven L.
REGISTRATION NUMBER: 37,642
REFERENCE/DOCKET NUMBER: UTSC:410/HYL
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
TELEX: 79-0924
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 46 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-021-625D-5

Query Match 100.0%; Score 103; DB 2; Length 46;
Best Local Similarity 100.0%; Pred. No. 2.5e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EMSPCSVTCGNGIOVRK 18
Db 29 EMSPCSVTCGNGIOVRK 46

RESULT 10
US-08-395-602A-4
Sequence 4, Application US/08395602A
Patent No. 5766899
GENERAL INFORMATION:
APPLICANT: Kuo, M. Tien
APPLICANT: Ding, Zhi-ming
TITLE OF INVENTION: Targeted Nucleic Acid Delivery into
TITLE OF INVENTION: Liver Cells
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: United States of America
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/395,602A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Highlander, Steven L.
REGISTRATION NUMBER: 37,642
REFERENCE/DOCKET NUMBER: UTSC:410/HYL
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
TELEX: 79-0924
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 126 amino acids
TYPE: amino acid
STRANDEDNESS: 1
TOPOLOGY: linear
US-08-395-602A-4

Query Match 100.0%; Score 103; DB 1; Length 126;
Best Local Similarity 100.0%; Pred. No. 7e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EWSPCSVTGNGIQRK 18
Db 109 EWSPCSVTGNGIQRK 126

RESULT 11
US-08-021-625D-4
Sequence 4, Application US/08021625D
Patent No. 5976851
GENERAL INFORMATION:
APPLICANT: Kuo, M. Tien
TITLE OF INVENTION: Targeted Nucleic Acid Delivery into
TITLE OF INVENTION: Liver Cells
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: United States of America
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/021,625D
FILING DATE: 16-FEB-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Highlander, Steven L.
REGISTRATION NUMBER: 37,642
REFERENCE/DOCKET NUMBER: UTSC:410/HYL
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
TELEX: 79-0924
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 126 amino acids
TYPE: amino acid
STRANDEDNESS: 1
TOPOLOGY: linear
US-08-021-625D-4

Query Match 100.0%; Score 103; DB 2; Length 126;
Best Local Similarity 100.0%; Pred. No. 7e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EWSPCSVTGNGIQRK 18
Db 109 EWSPCSVTGNGIQRK 126

RESULT 12
US-08-313-288B-18
Sequence 18, Application US/08313288B
Patent No. 5750502
GENERAL INFORMATION:
APPLICANT: Jesseil, Thomas M. and Avihu Klar
TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A
TITLE OF INVENTION: NOVEL SECRETED PROTEIN, F-SPONDIN
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/313,288B
FILING DATE: January 5, 1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 40028-A-PCT-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0526
TELEX:
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 412 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-313-288B-18

Query Match 100.0%; Score 103; DB 1; Length 412;
Best Local Similarity 100.0%; Pred. No. 2.4e-07;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EWSPCSVTGNGIQRK 18
Db 345 EWSPCSVTGNGIQRK 362

RESULT 13
US-08-760-797A-1
Sequence 1, Application US/08760797A
Patent No. 5928902
GENERAL INFORMATION:
APPLICANT: De Wilde, Michel
APPLICANT: Cohen, Joseph
TITLE OF INVENTION: Hybrid Protein Between CS
TITLE OF INVENTION: From Plasmodium and HBSAG
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA

ZIP: 19406
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/760,797A
FILING DATE: 04-DEC-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/442,612
FILING DATE: 17-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: Baumeister, Kirk
REGISTRATION NUMBER: 33,833
REFERENCE/DOCKET NUMBER: B45015-1C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-5096
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 423 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-760-797A-1

Query Match 100.0%; Score 103; DB 2; Length 423;
Best Local Similarity 100.0%; Pred. No. 2,4e-07;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EMSPCSTVCNGIQRK 18
|||
Db 139 EMSPCSTVCNGIQRK 156

RESULT 14
US-08-760-797A-3
Sequence 3, Application US/08760797A
Patent No. 5928902
GENERAL INFORMATION:
APPLICANT: De Wilde, Michel
APPLICANT: Cohen, Joseph
TITLE OF INVENTION: Hybrid Protein Between CS
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Smithkline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/760,797A
FILING DATE: 04-DEC-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/442,612
FILING DATE: 17-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: Baumeister, Kirk
REGISTRATION NUMBER: 33,833
REFERENCE/DOCKET NUMBER: B45015-1C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-5096

TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 424 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-760-797A-3

Query Match 100.0%; Score 103; DB 2; Length 424;
Best Local Similarity 100.0%; Pred. No. 2,4e-07;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EMSPCSTVCNGIQRK 18
|||
Db 136 EMSPCSTVCNGIQRK 153

RESULT 15
US-08-932-929B-1
Sequence 1, Application US/08932929B
Patent No. 6169171
GENERAL INFORMATION:
APPLICANT: De Wilde, Michel
APPLICANT: Cohen, Joseph
TITLE OF INVENTION: Hybrid Protein Between CS
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Smithkline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/932,929B
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/760,797
FILING DATE: 04-DEC-1996
APPLICATION NUMBER: 08/442,612
FILING DATE: 17-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: Baumeister, Kirk
REGISTRATION NUMBER: 33,833
REFERENCE/DOCKET NUMBER: B45015-1FWC2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-5096
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 424 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-932-929B-1

Query Match 100.0%; Score 103; DB 3; Length 424;
Best Local Similarity 100.0%; Pred. No. 2,4e-07;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EMSPCSTVCNGIQRK 18
|||
Db 140 EMSPCSTVCNGIQRK 157

Search completed: December 23, 2003, 15:37:39
Job time : 22.75 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 23, 2003, 15:33:48 ; Search time 36.75 Seconds
(without alignments)
91.474 Million cell updates/sec

Title: US-09-980-564-13

Perfect score: 103
Sequence: 1 EWSPTCVTCGNGICVRIK 18

Scoring table: BLOSUM62
Gap 10.0, Gapext 0.5

Searched: 696363 seqs, 186758610 residues

Total number of hits satisfying chosen parameters: 696363

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:.*
2: /cgn2_6/ptodata/1/pubpaa/PTCT_NEW_PUB.pep:.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:.*
6: /cgn2_6/ptodata/1/pubpaa/PTCT_PUBCOMB.pep:.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:.*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:.*
12: /cgn2_6/ptodata/1/pubpaa/US09C_NEW_PUB.pep:.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:.*
16: /cgn2_6/ptodata/1/pubpaa/US10C_NEW_PUB.pep:.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	103	100.0	18	US-10-098-238-1	Sequence 1, Appli
2	103	100.0	55	US-10-098-238-2	Sequence 2, Appli
3	103	100.0	55	US-10-098-238-3	Sequence 3, Appli
4	103	100.0	55	US-10-098-238-4	Sequence 4, Appli
5	103	100.0	55	US-10-098-238-5	Sequence 5, Appli
6	103	100.0	396	US-09-820-843A-31	Sequence 31, Appli
7	84	81.6	51	US-09-970-944-6	Sequence 46, Appli
8	84	81.6	51	US-10-190-115-92	Sequence 92, Appli
9	84	81.6	51	US-10-190-115-93	Sequence 93, Appli
10	84	81.6	51	US-10-190-115-94	Sequence 94, Appli
11	84	81.6	51	US-10-190-115-95	Sequence 95, Appli
12	84	81.6	51	US-09-981-151A-77	Sequence 77, Appli
13	83	80.6	18	US-10-098-238-7	Sequence 7, Appli
14	83	80.6	18	US-10-098-238-10	Sequence 10, Appli
15	83	80.6	20	US-10-098-238-14	Sequence 14, Appli

16	82	79.6	18	US-10-098-238-8	Sequence 8, Appli
17	82	79.6	18	US-10-098-238-9	Sequence 9, Appli
18	80	77.7	48	US-09-981-151A-78	Sequence 78, Appli
19	80	77.7	54	US-09-858-081-6	Sequence 6, Appli
20	80	77.7	54	US-09-858-081-6	Sequence 6, Appli
21	80	77.7	54	US-10-087-887-88	Sequence 88, Appli
22	80	77.7	54	US-10-087-887-90	Sequence 90, Appli
23	80	77.7	54	US-10-087-887-92	Sequence 92, Appli
24	80	77.7	54	US-10-014-070-9	Sequence 9, Appli
25	80	77.7	54	US-10-163-116-6	Sequence 6, Appli
26	80	77.7	54	US-10-163-547-19	Sequence 19, Appli
27	80	77.7	2150	US-09-321-987B-2	Sequence 2, Appli
28	80	77.7	2165	US-09-800-729-155	Sequence 155, Appli
29	77	74.8	16	US-10-098-238-6	Sequence 6, Appli
30	77	74.8	53	US-09-800-729-164	Sequence 164, Appli
31	77	74.8	525	US-09-918-171A-21	Sequence 21, Appli
32	77	74.8	525	US-09-978-295A-301	Sequence 301, Appli
33	77	74.8	525	US-09-978-697-301	Sequence 301, Appli
34	77	74.8	525	US-09-978-832A-301	Sequence 301, Appli
35	77	74.8	525	US-09-978-832A-301	Sequence 301, Appli
36	77	74.8	525	US-09-978-832A-301	Sequence 301, Appli
37	77	74.8	525	US-09-978-832A-301	Sequence 301, Appli
38	77	74.8	525	US-09-978-832A-301	Sequence 301, Appli
39	77	74.8	525	US-09-978-832A-301	Sequence 301, Appli
40	77	74.8	525	US-09-978-832A-301	Sequence 301, Appli
41	77	74.8	525	US-09-978-832A-301	Sequence 301, Appli
42	77	74.8	525	US-09-978-832A-301	Sequence 301, Appli
43	77	74.8	525	US-09-978-832A-301	Sequence 301, Appli
44	77	74.8	525	US-09-978-832A-301	Sequence 301, Appli
45	77	74.8	525	US-09-978-832A-301	Sequence 301, Appli

ALIGNMENTS

RESULT 1
US-10-098-238-1
; Sequence 1, Application US/10098238
; Publication No. US20030119733A1
GENERAL INFORMATION:
APPLICANT: Ceram, Carla
Inventor, Ute
Stimms, Photini
Nussenzweig, Victor
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
INHIBITING HEPATOCTE INVASION
BY MALARIAL SPOOROZOITES
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESSES:
ADDRESSER: Darby & Darby, P.C.
STREET: 805 Third Avenue
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5 inch, 360 KB storage
COMPUTER: COMPAQ PROLINEA MT 4/66
OPERATING SYSTEM: DOS 3.3
SOFTWARE: Word Perfect 5.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/098,238
FILING DATE: 14-Mar-2002
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/119,694
FILING DATE: 10-Sept-1993
ATTORNEY/AGENT INFORMATION:
NAME: Adda C. Gogoris, Esq.
REGISTRATION NUMBER: 29,714
REFERENCE/DOCKET NUMBER: 5986/17607-US1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 527-7700
TELEFAX: (212) 753-6237

TELEX: 236687
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 18 amino acids
TYPE: amino acid
TOPOLOGY: unknown
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-098-238-1

Query Match 100.0%; Score 103; DB 15; Length 18;
Best Local Similarity 100.0%; Pred. No. 2.5e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EWSPCSVTGNGIGVRIK 18
DB 1 EWSPCSVTGNGIGVRIK 18

RESULT 2
US-10-098-238-2
Sequence 2, Application US/10098238
Publication No. US20030119733A1
GENERAL INFORMATION:
APPLICANT: Cerami, Carla
Prevert, Ute
Simis, Photini
Nussenzweig, Victor
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
INHIBITING HEPATOCTYTE INVASION
BY MALARIAL SPOROZOITES
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: Darby & Darby, P.C.
STREET: 805 Third Avenue
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5, inch, 360 KB storage
COMPUTER: COMPAQ PROLINEA MT 4/66
OPERATING SYSTEM: DOS 3.3
SOFTWARE: Word Perfect 5.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/098,238
FILING DATE: 14-Mar-2002
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/119,694
FILING DATE: 10-Sept-1993
ATTORNEY/AGENT INFORMATION:
NAME: Adda C. Gogoris, Esq.
REGISTRATION NUMBER: 29,714
REFERENCE/DOCKET NUMBER: 5986/17607-US1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 527-7700
TELEFAX: (212) 753-6237
TELEX: 236687
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 55 amino acids
TYPE: amino acid
TOPOLOGY: unknown
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-098-238-2

Query Match 100.0%; Score 103; DB 15; Length 55;
Best Local Similarity 100.0%; Pred. No. 7.1e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EWSPCSVTGNGIGVRIK 18
DB 20 EWSPCSVTGNGIGVRIK 37

RESULT 3
US-10-098-238-3
Sequence 3, Application US/10098238
Publication No. US20030119733A1
GENERAL INFORMATION:
APPLICANT: Cerami, Carla
Prevert, Ute
Simis, Photini
Nussenzweig, Victor
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
INHIBITING HEPATOCTYTE INVASION
BY MALARIAL SPOROZOITES
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: Darby & Darby, P.C.
STREET: 805 Third Avenue
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5, inch, 360 KB storage
COMPUTER: COMPAQ PROLINEA MT 4/66
OPERATING SYSTEM: DOS 3.3
SOFTWARE: Word Perfect 5.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/098,238
FILING DATE: 14-Mar-2002
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/119,694
FILING DATE: 10-Sept-1993
ATTORNEY/AGENT INFORMATION:
NAME: Adda C. Gogoris, Esq.
REGISTRATION NUMBER: 29,714
REFERENCE/DOCKET NUMBER: 5986/17607-US1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 527-7700
TELEFAX: (212) 753-6237
TELEX: 236687
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 55 amino acids
TYPE: amino acid
TOPOLOGY: unknown
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-098-238-3

Query Match 100.0%; Score 103; DB 15; Length 55;
Best Local Similarity 100.0%; Pred. No. 7.1e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EWSPCSVTGNGIGVRIK 18
DB 20 EWSPCSVTGNGIGVRIK 37

RESULT 4
US-10-098-238-4
Sequence 4, Application US/10098238
Publication No. US20030119733A1
GENERAL INFORMATION:
APPLICANT: Cerami, Carla
Prevert, Ute
Simis, Photini
Nussenzweig, Victor
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
INHIBITING HEPATOCTYTE INVASION
BY MALARIAL SPOROZOITES
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: Darby & Darby, P.C.
STREET: 805 Third Avenue

CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5, inch, 360 KB storage
COMPUTER: COMPAQ PROLINEA MT 4/66
OPERATING SYSTEM: DOS 3.3
SOFTWARE: Word Perfect 5.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/098,238
FILING DATE: 14-Mar-2002
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/119,694
FILING DATE: 10-Sept-1993
ATTORNEY/AGENT INFORMATION:
NAME: Adda C. Gogoris, Esq.
REGISTRATION NUMBER: 29,714
REFERENCE/DOCKET NUMBER: 5986/17607-US1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 527-7700
TELEFAX: (212) 753-6237
TELEX: 236687
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 55 amino acids
TYPE: amino acid
TOPOLOGY: unknown
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-098-238-4

Query Match 100.0%; Score 103; DB 15; Length 55;
Best Local Similarity 100.0%; Pred. No. 7,1e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EWSPCSVTGNGIGVRIK 18
|||
Db 20 EWSPCSVTGNGIGVRIK 37

RESULT 5
US-10-098-238-5
Sequence 5, Application US/10098238
Publication No. US20030119733A1
GENERAL INFORMATION:
APPLICANT: Cerami, Carla
Frevert, Ute
Simms, Photini
Nussenzweig, Victor
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
INHIBITING HEPATOCYTE INVASION
BY MALARIAL SPROZOZOITES
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: Darby & Darby, P.C.
STREET: 805 Third Avenue
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5, inch, 360 KB storage
COMPUTER: COMPAQ PROLINEA MT 4/66
OPERATING SYSTEM: DOS 3.3
SOFTWARE: Word Perfect 5.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/098,238
FILING DATE: 14-Mar-2002
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/119,694
FILING DATE: 10-Sept-1993
ATTORNEY/AGENT INFORMATION:
NAME: Adda C. Gogoris, Esq.

REGISTRATION NUMBER: 29,714
REFERENCE/DOCKET NUMBER: 5986/17607-US1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 527-7700
TELEFAX: (212) 753-6237
TELEX: 236687
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 55 amino acids
TYPE: amino acid
TOPOLOGY: unknown
SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-10-098-238-5

Query Match 100.0%; Score 103; DB 15; Length 55;
Best Local Similarity 100.0%; Pred. No. 7,1e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EWSPCSVTGNGIGVRIK 18
|||
Db 20 EWSPCSVTGNGIGVRIK 37

RESULT 6
US-09-820-843A-31
Sequence 31, Application US/09820843A
Publication No. US20030039963A1
GENERAL INFORMATION:
APPLICANT: Council of Scientific and Industrial Research
TITLE OF INVENTION: A COMPUTATIONAL METHOD FOR THE IDENTIFICATION OF CANDIDATE PROTEIN
FILE REFERENCE: 063915
CURRENT APPLICATION NUMBER: US/09/820,843A
NUMBER OF SEQ ID NOS: 118
SOFTWARE: PatentIn version 3.0
SEQ ID NO 31
LENGTH: 396
TYPE: PRT
ORGANISM: Plasmodium falciparum
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Circumsporozoite (CS) protein
NAME/KEY: misc feature
OTHER INFORMATION: gi|4493889
US-09-820-843A-31

Query Match 100.0%; Score 103; DB 11; Length 396;
Best Local Similarity 100.0%; Pred. No. 4,6e-07;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EWSPCSVTGNGIGVRIK 18
|||
Db 329 EWSPCSVTGNGIGVRIK 346

RESULT 7
US-09-970-944-46
Sequence 46, Application US/09970944
Publication No. US20030204052A1
GENERAL INFORMATION:
APPLICANT: Herrman, John L.
Rastelli, Luca
APPLICANT: Shinkets, Richard A
TITLE OF INVENTION: No. US20030204052A1el Proteins and Nucleic Acids Encoding Same an
FILE REFERENCE: 21402-138
CURRENT APPLICATION NUMBER: US/09/970,944
NUMBER OF FILING DATE: 2002-05-02
PRIOR APPLICATION NUMBER: 60/237,862
FILING DATE: 2000-10-04
NUMBER OF SEQ ID NOS: 62
SOFTWARE: PatentIn Ver. 2.1

ORGANISM: Homo sapiens
US-10-190-115-93

Query Match 81.6%; Score 84; DB 12; Length 51;
Best Local Similarity 72.2%; Pred. No. 3.7e-05;
Matches 13; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EWSPCSVTCGNGIQVRIK 18
|||||||:
DB 6 EWSPCSVTCGGVQTRTR 23

RESULT 10
US-10-190-115-94

Sequence 94, Application US/10190115
Publication No. US20030207394A1
GENERAL INFORMATION:

APPLICANT: Alsobrook, John P. II
APPLICANT: Boldog, Ferenc L.
APPLICANT: Burgess, Catherine E.
APPLICANT: Casman, Stacie J.
APPLICANT: Grosse, William M.
APPLICANT: Gusev, Vladimir Y.
APPLICANT: Ji, Weizhen
APPLICANT: Lepley, Denise M.
APPLICANT: Liu, Xiaohong
APPLICANT: Mezik, Amanda J.
APPLICANT: Padigaru, Muralidhara
APPLICANT: Patturajan, Meera
APPLICANT: Rastelli, Luca
APPLICANT: Shen, Lei
APPLICANT: Shenoy, Suresh G.
APPLICANT: Shinkets, Richard A.
APPLICANT: Spaderna, Steven K.
APPLICANT: Spytek, Kimberly A.
APPLICANT: Szekeres, Edward S. Jr.
APPLICANT: Taupier, Raymond J. Jr.
APPLICANT: Tchernev, Velizar T.
APPLICANT: Zernusen, Bryan D.
APPLICANT: Voss, Edward Z.
TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
FILE REFERENCE: 21402-050 CIP
CURRENT FILING DATE: 2003-02-10
PRIOR FILING DATE: 2003-02-10
PRIOR APPLICATION NUMBER: 60/303,168
PRIOR FILING DATE: 2001-07-05
PRIOR APPLICATION NUMBER: 60/368,996
PRIOR FILING DATE: 2002-04-01
PRIOR APPLICATION NUMBER: 60/386,816
PRIOR FILING DATE: 2002-06-07
PRIOR APPLICATION NUMBER: 60/215,854
PRIOR FILING DATE: 2000-07-03
PRIOR APPLICATION NUMBER: 60/215,856
PRIOR FILING DATE: 2000-07-03
PRIOR APPLICATION NUMBER: 60/215,902
PRIOR FILING DATE: 2000-07-03
PRIOR APPLICATION NUMBER: 60/216,585,
PRIOR FILING DATE: 2000-07-07
PRIOR APPLICATION NUMBER: 60/216,586
PRIOR FILING DATE: 2001-07-07
PRIOR APPLICATION NUMBER: 60/216,722
PRIOR FILING DATE: 2000-07-07
PRIOR APPLICATION NUMBER: 60/218,622
PRIOR FILING DATE: 2000-07-17
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 136
SOFTWARE: CuroSeqList version 0.1
SEQ ID NO 94
TYPE: PRT
LENGTH: 51
ORGANISM: Homo sapiens
US-10-190-115-94

Query Match 81.6%; Score 84; DB 12; Length 51;
Best Local Similarity 72.2%; Pred. No. 3.7e-05;
Matches 13; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EWSPCSVTCGNGIQVRIK 18
|||||||:
DB 6 EWSPCSVTCGGVQTRTR 23

RESULT 11
US-10-190-115-95

Sequence 95, Application US/10190115
Publication No. US20030207394A1
GENERAL INFORMATION:

APPLICANT: Alsobrook, John P. II
APPLICANT: Boldog, Ferenc L.
APPLICANT: Burgess, Catherine E.
APPLICANT: Casman, Stacie J.
APPLICANT: Grosse, William M.
APPLICANT: Gusev, Vladimir Y.
APPLICANT: Ji, Weizhen
APPLICANT: Lepley, Denise M.
APPLICANT: Liu, Xiaohong
APPLICANT: Mezik, Amanda J.
APPLICANT: Padigaru, Muralidhara
APPLICANT: Patturajan, Meera
APPLICANT: Rastelli, Luca
APPLICANT: Shen, Lei
APPLICANT: Shenoy, Suresh G.
APPLICANT: Shinkets, Richard A.
APPLICANT: Spaderna, Steven K.
APPLICANT: Spytek, Kimberly A.
APPLICANT: Szekeres, Edward S. Jr.
APPLICANT: Taupier, Raymond J. Jr.
APPLICANT: Tchernev, Velizar T.
APPLICANT: Zernusen, Bryan D.
APPLICANT: Voss, Edward Z.
TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
FILE REFERENCE: 21402-050 CIP
CURRENT FILING DATE: 2003-02-10
PRIOR FILING DATE: 2003-02-10
PRIOR APPLICATION NUMBER: 60/303,168
PRIOR FILING DATE: 2001-07-05
PRIOR APPLICATION NUMBER: 60/368,996
PRIOR FILING DATE: 2002-04-01
PRIOR APPLICATION NUMBER: 60/386,816
PRIOR FILING DATE: 2002-06-07
PRIOR APPLICATION NUMBER: 60/215,854
PRIOR FILING DATE: 2000-07-03
PRIOR APPLICATION NUMBER: 60/215,856
PRIOR FILING DATE: 2000-07-03
PRIOR APPLICATION NUMBER: 60/215,902
PRIOR FILING DATE: 2000-07-03
PRIOR APPLICATION NUMBER: 60/216,585,
PRIOR FILING DATE: 2000-07-07
PRIOR APPLICATION NUMBER: 60/216,586
PRIOR FILING DATE: 2001-07-07
PRIOR APPLICATION NUMBER: 60/216,722
PRIOR FILING DATE: 2000-07-07
PRIOR APPLICATION NUMBER: 60/218,622
PRIOR FILING DATE: 2000-07-17
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 136
SOFTWARE: CuroSeqList version 0.1
SEQ ID NO 95
TYPE: PRT
LENGTH: 51
ORGANISM: Homo sapiens
US-10-190-115-95

Query Match 81.6%; Score 84; DB 12; Length 51;
Best Local Similarity 72.2%; Pred. No. 3.7e-05;
Matches 13; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EMSPCSVTCGNGIOVRK 18
|||:|:|:
Db 6 EMSPCSVTCGGVQTRR 23

RESULT 12

US-09-981-151A-77
; Sequence 77, Application US/09981151A
; Publication No. US20030212256A1
; GENERAL INFORMATION:
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Gerlach, Valerie
; APPLICANT: MacDougall, John R
; APPLICANT: Malyankar, Muriel M
; APPLICANT: Smithson, Glenda
; APPLICANT: Millet, Isabelle
; APPLICANT: Peyman, John A
; APPLICANT: Stone, David J
; APPLICANT: Gunther, Erik
; APPLICANT: Ellerman, Karen
; APPLICANT: Shinkets, Richard A
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Guo, Xiaojia
; APPLICANT: Paturajan, Meera
; APPLICANT: Taulier Jr, Raymond J
; APPLICANT: Burgess, Catherine E
; APPLICANT: Zerhusen, Bryan D
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Gangolli, Esha A
; APPLICANT: Fernandes, Elma R
; APPLICANT: Gorman, Linda
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-168
; CURRENT APPLICATION NUMBER: US/09/981,151A
; CURRENT FILING DATE: 2001-10-16
; PRIOR APPLICATION NUMBER: 60/241,040
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/241,058
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/241,063
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/241,243
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/242,152
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/242,482
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/242,611
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/242,612
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/242,880
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: 60/242,881
; PRIOR FILING DATE: 2000-10-24
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 160
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 77
; LENGTH: 51
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Thrombospondin
; OTHER INFORMATION: type 1 Consensus Sequence
US-09-981-151A-77

Query Match 81.6%; Score 84; DB 12; Length 51;
Best Local Similarity 72.2%; Pred. No. 3.7e-05;
Matches 13; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EMSPCSVTCGNGIOVRK 18
|||:|:|:
Db 6 EMSPCSVTCGGVQTRR 23

RESULT 13

US-10-098-238-7
; Sequence 7, Application US/10098238
; Publication No. US20030119733A1
; GENERAL INFORMATION:
; APPLICANT: Cerami, Carla
; Frevert, Ute
; Simlis, Photini
; Nussenzweig, Victor
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; INHIBITING HEPATOCTE INVASION
; BY MALARIAL SPOROZOITES
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Darby & Darby, P.C.
; STREET: 805 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5, inch, 360 KB storage
; OPERATING SYSTEM: DOS 3.3
; SOFTWARE: Word Perfect 5.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/098,238
; FILING DATE: 14-Mar-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/119,694
; FILING DATE: 10-Sept-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Adda C. Gogoris, Esq.
; REGISTRATION NUMBER: 29,714
; REFERENCE/DOCKET NUMBER: 5986/17607-US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 527-7700
; TELEFAX: (212) 753-6237
; TELEX: 236687
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 18 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; US-10-098-238-7

Query Match 80.6%; Score 83; DB 15; Length 18;
Best Local Similarity 77.8%; Pred. No. 1.9e-05;
Matches 14; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EMSPCSVTCGNGIOVRK 18
|||:|:|:
Db 1 EMSPCSVTCGGIRARK 18

RESULT 14

US-10-098-238-10
; Sequence 10, Application US/10098238
; Publication No. US20030119733A1
; GENERAL INFORMATION:
; APPLICANT: Cerami, Carla
; Frevert, Ute
; Simlis, Photini
; Nussenzweig, Victor
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; INHIBITING HEPATOCTE INVASION
; BY MALARIAL SPOROZOITES

NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: Darby & Darby, P.C.
STREET: 805 Third Avenue
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5, inch, 360 KB storage
COMPUTER: COMPAQ PROLINEA MT 4/66
OPERATING SYSTEM: DOS 3.3
SOFTWARE: Word Perfect 5.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/098,238
FILING DATE: 14-Mar-2002
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/119,694
FILING DATE: 10-Sept-1993
ATTORNEY/AGENT INFORMATION:
NAME: Adda C. Gogoris, Esq.
REGISTRATION NUMBER: 29,714
REFERENCE/DOCKET NUMBER: 5986/17607-US1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 527-7700
TELEFAX: (212) 753-6237
TELEX: 236687
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 18 amino acids
TYPE: amino acid
TOPOLOGY: unknown
SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-10-098-238-10
Query Match 80.6%; Score 83; DB 15; Length 18;
Best Local Similarity 77.8%; Pred. No. 1.9e-05;
Matches 14; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 EMSPCVTGCGNGIQVRIK 18
Db 1 EMSPCVTGCGNGIRARRK 18
RESULT 15
US-10-098-238-14
Sequence 14, Application US/10098238
Publication No. US20030119733A1
GENERAL INFORMATION:
APPLICANT: Ceram!, Carla
Frevert, Ute
Simmls, Photini
Nussenzeig, Victor
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
INHIBITING HEPATOCYTE INVASION
BY MALARIAL SPOROZOITES
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: Darby & Darby, P.C.
STREET: 805 Third Avenue
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5, inch, 360 KB storage
COMPUTER: COMPAQ PROLINEA MT 4/66
OPERATING SYSTEM: DOS 3.3
SOFTWARE: Word Perfect 5.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/098,238
FILING DATE: 14-Mar-2002
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/119,694
FILING DATE: 10-Sept-1993
ATTORNEY/AGENT INFORMATION:
NAME: Adda C. Gogoris, Esq.
REGISTRATION NUMBER: 29,714
REFERENCE/DOCKET NUMBER: 5986/17607-US1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 527-7700
TELEFAX: (212) 753-6237
TELEX: 236687
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
TOPOLOGY: unknown
SEQUENCE DESCRIPTION: SEQ ID NO: 14:
US-10-098-238-14
Query Match 80.6%; Score 83; DB 15; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.1e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 PCSVTGCGNGIQVRIK 18
Db 1 PCSVTGCGNGIQVRIK 15
Search completed: December 23, 2003, 15:38:41
Job time : 37.75 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compen Ltd.

OM protein - protein search, using sw model

Run on: December 23, 2003, 15:33:08 / Search time 20.25 Seconds

(without alignments)
85.483 Million cell updates/sec

Title: US-09-980-564-13

Perfect score: 103

Sequence: 1 EMSPCSVTCGNGICQVRIK 18

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

- Listing first 45 summaries

Database :

1: PIR_76:*

2: PIR1:*

3: PIR2:*

4: PIR3:*

5: PIR4:*

6: PIR5:*

7: PIR6:*

8: PIR7:*

9: PIR8:*

10: PIR9:*

11: PIR10:*

12: PIR11:*

13: PIR12:*

14: PIR13:*

15: PIR14:*

16: PIR15:*

17: PIR16:*

18: PIR17:*

19: PIR18:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	103	100.0	388	A39756	circumsporozoite p
2	103	100.0	405	S05428	circumsporozoite p
3	103	100.0	412	OZQAF	circumsporozoite p
4	103	100.0	424	A54533	circumsporozoite p
5	103	100.0	442	A54529	circumsporozoite p
6	87	84.5	351	OZQAK	circumsporozoite p
7	87	84.5	363	OZQAK	circumsporozoite p
8	83	80.6	429	A54504	circumsporozoite p
9	83	80.6	485	A60610	circumsporozoite p
10	82	79.6	378	OZQAB	circumsporozoite p
11	82	79.6	378	OZQAL	circumsporozoite p
12	82	79.6	398	OZQAS	circumsporozoite p
13	82	79.6	401	OZQAC	circumsporozoite p
14	82	79.6	419	OZQAM	circumsporozoite p
15	80	77.7	2165	T21371	hypothetical prote
16	77	74.8	343	A29319	circumsporozoite p
17	77	74.8	367	A23068	circumsporozoite p
18	77	74.8	378	OZQAV	circumsporozoite p
19	77	74.8	386	A48571	circumsporozoite p
20	77	74.8	387	D41156	circumsporozoite p
21	77	74.8	387	D41156	circumsporozoite p
22	77	74.8	395	A41156	circumsporozoite p
23	76	73.8	264	A44969	circumsporozoite p
24	76	73.8	264	OZQOM	circumsporozoite p
25	75	72.8	388	JG6164	circumsporozoite p
26	75	72.8	1444	T18856	angiogenesis inhib
27	74	71.8	332	OZQMB	circumsporozoite p
28	74	71.8	348	OZQKB	circumsporozoite p
29	73	70.9	559	S04531	thrombospondin-rel

30	73	70.9	574	2	A46283	sporozoite surface
31	73	70.9	1170	2	A40558	thrombospondin 1 p
32	72.5	70.4	1572	2	T00027	brain-specific ang
33	72	69.9	984	2	T00326	hypothetical prote
34	72	69.9	1074	2	UC5928	hemaphorin F precu
35	72	69.9	1522	2	T00028	brain-specific ang
36	72	69.9	2098	2	T18397	protein CTRP - mal
37	71	68.9	1558	2	C89114	protein C37C3.6a l
38	71	68.9	2167	2	T34395	hypothetical prote
39	70	68.0	803	2	A47723	F-spondin precuro
40	70	68.0	1205	2	T18517	procollagen N-endo
41	69	67.0	805	2	T34212	hypothetical prote
42	69	67.0	807	2	A38152	F-spondin - rat
43	69	67.0	1170	2	TSHUP1	thrombospondin 1 p
44	67	65.0	437	2	S05478	propeptin - mouse
45	67	65.0	1172	1	TSHUP2	thrombospondin 2 p

ALIGNMENTS

RESULT 1

A39756

Circumsporozoite protein - Plasmodium reichenowi

C:Species: Plasmodium reichenowi

C:Date: 14-Feb-1992 #sequence_revision 14-Feb-1992 #text_change 20-Aug-1999

C:Accession: A39756

R:Blal, A.A.; Goldman, I.F.

U. Biol. Chem. 266, 6686-6689, 1991

A:Title: Circumsporozoite protein gene from Plasmodium reichenowi, a chimpanzee malaria

A:Reference number: A39756; MUID:91201303; PMID:2016283

A:Accession: A39756

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-388 <LAI>

A:Cross-references: GB:M60972; NID:9160228; PIDN:AAA29561.1; PID:9160229

C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology

F:312-366/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match	Score	DB 2;	Length	388;
Best Local Similarity	100.0%;	Pred. No. 9.1e-08;		
Matches	18;	Conservative	0;	Mismatches 0;
				Indels 0;
				Gaps 0;

QY 1 EMSPCSVTCGNGICQVRIK 18

Db 321 EMSPCSVTCGNGICQVRIK 338

RESULT 2

S05428

Circumsporozoite protein - malaria parasite (Plasmodium falciparum) (isolate NF54)

C:Species: Plasmodium falciparum

C:Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 09-Jun-2000

C:Accession: S05428; A45527; I60657

R:Campbell, J.R.

Nucleic Acids Res. 17, 5854, 1989

A:Title: DNA sequence of the gene encoding a Plasmodium falciparum malaria candidate vac

A:Reference number: S05428; MUID:89345189; PMID:266895

A:Accession: S05428

A:Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-405 <CAM>

A:Cross-references: EMBL:X15363

R:Caperez, P.; Gentz, R.; Watile, H.; Pink, J.R.; Sinigaglia, F.

Mol. Biochem. Parasitol. 35, 185-190, 1989

A:Title: The circumsporozoite protein gene from NF54, a Plasmodium falciparum isolate us

A:Reference number: A45527; MUID:89345189; PMID:2671723

A:Accession: A45527

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-405 <CAS>

A:Cross-references: GB:M22982; GB:J04650; NID:9160168; PIDN:AAA29527.1; PID:9160169

R:Lockyer, M.J.; Marsh, K.; Newbold, C.I.

Mol. Biochem. Parasitol. 37, 275-280, 1989
 A>Title: Wild isolates of Plasmodium falciparum show extensive polymorphism in T cell ep
 A:Reference number: A60657; PMID:90114334; PMID:2481827
 A:Accession: 160657
 A>Status: preliminary; not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 319-336,354-373 <LOC>
 C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
 F:329-383/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match 100.0%; Score 103; DB 2; Length 405;
 Best Local Similarity 100.0%; Pred. No. 9.4e-08;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 EWSPCSVTGNGIGVRIK 18
 |||
 Db 338 EWSPCSVTGNGIGVRIK 355

RESULT 3
 OZ20AF
 circumsporozoite protein - malaria parasite (Plasmodium falciparum) (isolate INTM22)
 C:Species: Plasmodium falciparum
 C>Date: 15-Nov-1984 #sequence_revision 15-Nov-1984 #text_change 09-Jun-2000
 C:Accession: A03388
 R:Dame, J.B.; Williams, J.L.; McCutchan, T.F.; Weber, J.L.; Wirtz, R.A.; Hochmeyer, W.T.
 Science 225, 593-599, 1984
 A>Title: Structure of the gene encoding the immunodominant surface antigen on the sporoz
 A:Reference number: A03388; PMID:84250215; PMID:6204383
 A:Accession: A03388
 A:Molecule type: DNA
 A:Residues: 1-412 <DM>
 A:Cross-references: GB:K02194; NID:g160160; PIDN:AAA29524.1; PID:g160161
 A:Experimental source: clone 7G8
 C:Comment: Residues 1-16 are the probable signal sequence.
 C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
 F:336-390/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match 100.0%; Score 103; DB 1; Length 412;
 Best Local Similarity 100.0%; Pred. No. 9.5e-08;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 EWSPCSVTGNGIGVRIK 18
 |||
 Db 345 EWSPCSVTGNGIGVRIK 362

RESULT 4
 A54533
 circumsporozoite protein - malaria parasite (Plasmodium falciparum) (strain T4, Thailand
 C:Species: Plasmodium falciparum
 C>Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 09-Jun-2000
 C:Accession: A54533
 R:del Portillo, H.A.; Nussenzweig, R.S.; Enea, V.
 Mol. Biochem. Parasitol. 24, 289-294, 1987
 A>Title: Circumsporozoite gene of a Plasmodium falciparum strain from Thailand.
 A:Reference number: A54533; PMID:87315205; PMID:3306373
 A:Accession: A54533
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-424
 A:Cross-references: GB:M9752; NID:g160216; PIDN:AAA29555.1; PID:g160217
 C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
 F:348-402/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match 100.0%; Score 103; DB 2; Length 424;
 Best Local Similarity 100.0%; Pred. No. 9.8e-08;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 EWSPCSVTGNGIGVRIK 18
 |||
 Db 357 EWSPCSVTGNGIGVRIK 374

RESULT 5
 A54529
 circumsporozoite protein - malaria parasite (Plasmodium falciparum) (strain Wellcome)
 C:Species: Plasmodium falciparum
 C>Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 09-Jun-2000
 C:Accession: A54529
 R:Lockyer, M.J.; Schwarz, R.T.
 Mol. Biochem. Parasitol. 22, 101-108, 1987
 A>Title: Strain variation in the circumsporozoite protein gene of Plasmodium falciparum.
 A:Reference number: A54529; PMID:87115616; PMID:3543671
 A:Accession: A54529
 A>Status: preliminary; not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 1-442 <LOC>
 A:Cross-references: GB:M15505; NID:g160214; PIDN:AAA29554.1; PID:g160215
 C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
 F:366-420/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match 100.0%; Score 103; DB 2; Length 442;
 Best Local Similarity 100.0%; Pred. No. 1e-07;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 EWSPCSVTGNGIGVRIK 18
 |||
 Db 375 EWSPCSVTGNGIGVRIK 392

RESULT 6
 OZ20KV
 circumsporozoite protein precursor - Plasmodium knowlesi (strain Nuri)
 N:Alternate names: sporozoite surface protein
 C:Species: Plasmodium knowlesi
 C>Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 16-Jul-1999
 C:Accession: A26253
 R/Sharma, S.; Svec, P.; Mitchell, G.H.; Godson, G.N.
 Science 229, 779-782, 1985
 A>Title: Diversity of circumsporozoite antigen genes from two strains of the malarial par
 A:Reference number: A26253; PMID:85272582; PMID:4023712
 A:Accession: A26253
 A:Molecule type: DNA
 A:Residues: 1-351 <SHA>
 A:Cross-references: GB:M11031; NID:g160197; PIDN:AAA29540.1; PID:g160198
 C:Comment: There are three distinct regions in the mature circumsporozoite protein, the e
 obic membrane-anchoring sequence.
 C:Comment: There are 14 tandem copies of the 9-residue repeat E-Q-P-A-A-G-A-G/R-G (plus t
 C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
 F:1-19/Domain: signal sequence #status predicted <SIG>
 F:20-351/Product: circumsporozoite protein #status predicted <MAT>
 F:98-223/Region: 9-residue repeats
 F:224-241/Region: 9-residue repeats
 F:276-329/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match 84.5%; Score 87; DB 1; Length 351;
 Best Local Similarity 72.2%; Pred. No. 2e-05;
 Matches 13; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

1 EWSPCSVTGNGIGVRIK 18
 |||
 Db 285 EWTPCSVTGNGVIRRK 302

RESULT 7
 OZ20AK
 circumsporozoite protein - Plasmodium knowlesi (strain H)
 C:Species: Plasmodium knowlesi
 C>Date: 03-Aug-1984 #sequence_revision 03-Aug-1984 #text_change 16-Jul-1999
 C:Accession: A90841; A93115; A03389
 R:ozaki, L.S.; Svec, P.; Nussenzweig, R.S.; Nussenzweig, V.; Godson, G.N.
 Cell 34, 815-822, 1983

A>Title: Structure of the Plasmodium knowlesi gene coding for the circumsporozoite prote
 A:Reference number: A90841; MUID:84026486; PMID:6313209
 A:Accession: A90841
 A:Molecule type: DNA
 A:Residues: 1-363 <OZA>
 A:Cross-references: GB:K00822; NID:9160195; PIDN:AAA19699.1; PID:9160196
 R:Gordon, G.N.; Ellis, J.; Svec, P.; Schlesinger, D.H.; Nussenzweig, V.
 Nature 305, 29-33, 1983
 A>Title: Identification and chemical synthesis of a tandemly repeated immunogenic region
 A:Reference number: A93315; MUID:83297689; PMID:6193427
 A:Accession: A93315
 A:Molecule type: mRNA
 A:Residues: 107-208 <GOD>
 A:Cross-references: GB:K00772
 C:Comment: At least 12 copies of a 12-residue repeating unit occur in this surface prote
 e mosquito to the vertebrate host.
 C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
 F:288-341/Domain: thrombospondin type 1 repeat homology <THR1>
 F:288-341/Domain: thrombospondin type 1 repeat homology <THR1>
 Query Match 84.5%; Score 87; DB 1; Length 363;
 Best Local Similarity 72.2%; Pred. No. 2e-05;
 Matches 13; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
 QY 1 EMSPCSVTCGNGIOVRK 18
 |||:|||||:|||||:|
 Db 297 EWTPCSVTCGNGIRARRK 314
 |||:|||||:|||||:|
 RESULT 8
 A54504
 C:circumsporozoite protein - Plasmodium malariae (strain Uganda 1-CDC)
 C:Species: Plasmodium malariae
 C:Date: 06-Oct-1994 #sequence_revision 04-Nov-1994 #text_change 20-Aug-1999
 C:Accession: A54504
 R:La, A.A.; de la Cruz, V.F.; Campbell, G.H.; Procell, P.M.; Collins, W.E.; McCutchan,
 M.O. Biochem. Parasitol. 30, 291-294, 1988
 A>Title: Structure of the circumsporozoite gene of Plasmodium malariae.
 A:Reference number: A54504; MUID:89040027; PMID:3054537
 A:Accession: A54504
 A:Molecule type: DNA
 A:Residues: 1-429 <LAL>
 A:Cross-references: GB:J03992; NID:9160220; PIDN:AAA29557.1; PID:9160221
 C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
 C:Keywords: tandem repeat
 F:354-407/Domain: thrombospondin type 1 repeat homology <THR1>
 Query Match 80.6%; Score 83; DB 2; Length 429;
 Best Local Similarity 77.8%; Pred. No. 9.1e-05;
 Matches 14; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 1 EMSPCSVTCGNGIOVRK 18
 |||:|||||:|||||:|
 Db 363 EMSPCSVTCGNGIRARRK 380
 |||:|||||:|||||:|
 RESULT 9
 A60610
 C:circumsporozoite protein precursor - Plasmodium brasiliannu
 C:Species: Plasmodium brasiliannu
 C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 20-Aug-1999
 C:Accession: A60610; A28615
 R:di Giovanni, L.; Cochran, A.H.; Enea, V.
 Exp. Parasitol. 70, 373-381, 1990
 A>Title: On the evolutionary history of the circumsporozoite protein in plasmodia.
 A:Reference number: A60610; MUID:90214818; PMID:2323391
 A:Accession: A60610
 A:Status: not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 1-485 <DIA>
 A:Experimental source: strain Colombian
 R:La, A.A.; de la Cruz, V.F.; Collins, W.E.; Campbell, G.H.; Procell, P.M.; McCutchan,
 J. Biol. Chem. 263, 5495-5498, 1988
 A>Title: Circumsporozoite protein gene from Plasmodium brasiliannu. Animal reservoirs fd

A:Reference number: A28615; MUID:88186854; PMID:3128542
 A:Accession: A28615
 A:Molecule type: DNA
 A:Residues: 93-485 <LAL>
 A:Cross-references: GB:J03203; NID:9160212; PIDN:AAA29553.1; PID:9160213
 C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
 C:Keywords: sporozoite; surface antigen; tandem repeat
 F:1-20/Domain: signal sequence #status predicted <SIG>
 F:21-485/Product: circumsporozoite protein #status predicted <MAT>
 F:114-369/Region: 4-residue repeats
 F:410-463/Domain: thrombospondin type 1 repeat homology <THR1>
 Query Match 80.6%; Score 83; DB 2; Length 485;
 Best Local Similarity 77.2%; Pred. No. 0.0001;
 Matches 14; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 1 EMSPCSVTCGNGIOVRK 18
 |||:|||||:|||||:|
 Db 419 EMSPCSVTCGNGIRARRK 436
 |||:|||||:|||||:|
 RESULT 10
 OZZOAB
 C:circumsporozoite protein precursor - Plasmodium cynomolgi (strain BeroK)
 N:Alternate names: major sporozoite surface antigen
 C:Species: Plasmodium cynomolgi
 C:Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 07-Nov-1997
 C:Accession: D26255
 R:Galinski, M.R.; Arnot, D.E.; Cochran, A.H.; Barnwell, J.W.; Nussenzweig, R.S.; Enea,
 Cell 48, 311-319, 1987
 A>Title: The circumsporozoite gene of the Plasmodium cynomolgi complex.
 A:Reference number: A90889; MUID:87102878; PMID:3802196
 A:Accession: D26255
 A:Molecule type: DNA
 A:Residues: 1-378 <GAL>
 C:Comment: There are three distinct regions in the mature circumsporozoite protein, the
 obic membrane-anchoring sequence.
 C:Comment: There are 10 tandem copies of a 9-residue repeat (preceded by a 6-residue inc
 C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
 C:Keywords: sporozoite; surface antigen; tandem repeat
 F:1-19/Domain: signal sequence #status predicted <SIG>
 F:20-378/Product: circumsporozoite protein #status predicted <MAT>
 F:97-192/Region: 9-residue repeats
 F:193-268/Region: 16-residue repeats
 F:303-356/Domain: thrombospondin type 1 repeat homology <THR1>
 Query Match 79.6%; Score 82; DB 1; Length 378;
 Best Local Similarity 72.2%; Pred. No. 0.00011;
 Matches 13; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 QY 1 EMSPCSVTCGNGIOVRK 18
 |||:|||||:|||||:|
 Db 312 EMSPCSVTCGNGIRARRK 329
 |||:|||||:|||||:|
 RESULT 11
 OZZOAL
 C:circumsporozoite protein precursor - Plasmodium cynomolgi (strain London)
 N:Alternate names: major sporozoite surface antigen
 C:Species: Plasmodium cynomolgi
 C:Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 07-Nov-1997
 C:Accession: A26255
 R:Galinski, M.R.; Arnot, D.E.; Cochran, A.H.; Barnwell, J.W.; Nussenzweig, R.S.; Enea,
 Cell 48, 311-319, 1987
 A>Title: The circumsporozoite gene of the Plasmodium cynomolgi complex.
 A:Reference number: A90889; MUID:87102878; PMID:3802196
 A:Accession: A26255
 A:Molecule type: DNA
 A:Residues: 1-378 <GAL>
 C:Comment: There are three distinct regions in the mature circumsporozoite protein, the
 obic membrane-anchoring sequence.
 C:Comment: There are 19 tandem copies of a 6-residue repeat and 6 copies of an 11-residu
 C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology

C:Keywords: sporozoite; surface antigen; tandem repeat
 F:1-19/Domain: signal sequence #status predicted <SIG>
 F:20-378/Product: circumsporozoite protein #status predicted <MAT>
 F:98-211/Region: 6-residue repeats
 F:212-277/Region: 11-residue repeats
 F:303-356/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match 79.6%; Score 82; DB 1; Length 378;
 Best Local Similarity 72.2%; Pred. No. 0.00011;
 Matches 13; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EWSPCSVTGNGIQVRIK 18
 |||||
 Db 312 EWSPCSVTGKGVRRRK 329

RESULT 12

OZZQNC
 circumsporozoite protein precursor - Plasmodium cynomolgi (strain Ceylon)
 N:Alternate names: major sporozoite surface antigen
 C:Species: Plasmodium cynomolgi

C>Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 07-Nov-1997
 C/Accession: C26255
 R:Galinski, M.R.; Arnot, D.E.; Cochran, A.H.; Barnwell, J.W.; Nussenzweig, R.S.; Enea, Cell 48, 311-319, 1987
 A:Title: The circumsporozoite gene of the Plasmodium cynomolgi complex.
 A:Reference number: A90889; MUID:87102878; PMID:3802196
 A:Accession: C26255

A:Molecule type: DNA
 A:Residues: 1-398 <GAL>
 C:Comment: There are three distinct regions in the mature circumsporozoite protein, the obic membrane-anchoring sequence.

C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
 C:Keywords: sporozoite; surface antigen; tandem repeat
 F:1-19/Domain: signal sequence #status predicted <SIG>
 F:20-398/Product: circumsporozoite protein #status predicted <MAT>
 F:97-240/Region: 9-residue repeats
 F:241-291/Region: 17-residue repeats
 F:323-376/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match 79.6%; Score 82; DB 1; Length 398;
 Best Local Similarity 72.2%; Pred. No. 0.00012;
 Matches 13; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EWSPCSVTGNGIQVRIK 18
 |||||
 Db 332 EWSPCSVTGKGVRRRK 349

RESULT 13

OZZQNC
 circumsporozoite protein precursor - Plasmodium cynomolgi (strain Gombak)
 N:Alternate names: major sporozoite surface antigen
 C:Species: Plasmodium cynomolgi

C>Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 07-Nov-1997
 C/Accession: E26255
 R:Galinski, M.R.; Arnot, D.E.; Cochran, A.H.; Barnwell, J.W.; Nussenzweig, R.S.; Enea, Cell 48, 311-319, 1987
 A:Title: The circumsporozoite gene of the Plasmodium cynomolgi complex.
 A:Reference number: A90889; MUID:87102878; PMID:3802196
 A:Accession: E26255

A:Molecule type: DNA
 A:Residues: 1-401 <GAL>
 C:Comment: There are three distinct regions in the mature circumsporozoite protein, the obic membrane-anchoring sequence.

C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
 C:Keywords: sporozoite; surface antigen; tandem repeat
 F:1-19/Domain: signal sequence #status predicted <SIG>
 F:20-401/Product: circumsporozoite protein #status predicted <MAT>
 F:98-278/Region: 11-residue repeats
 F:326-379/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match 79.6%; Score 82; DB 1; Length 401;
 Best Local Similarity 72.2%; Pred. No. 0.00012;
 Matches 13; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EWSPCSVTGNGIQVRIK 18
 |||||
 Db 335 EWSPCSVTGKGVRRRK 352

RESULT 14

OZZQNC
 circumsporozoite protein precursor - Plasmodium cynomolgi (strain Mulligan/NIH)
 N:Alternate names: major sporozoite surface antigen
 C:Species: Plasmodium cynomolgi

C>Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 07-Nov-1997
 C/Accession: B26255
 R:Galinski, M.R.; Arnot, D.E.; Cochran, A.H.; Barnwell, J.W.; Nussenzweig, R.S.; Enea, Cell 48, 311-319, 1987
 A:Title: The circumsporozoite gene of the Plasmodium cynomolgi complex.
 A:Reference number: A90889; MUID:87102878; PMID:3802196
 A:Accession: B26255

A:Molecule type: DNA
 A:Residues: 1-419 <GAL>
 C:Comment: There are three distinct regions in the mature circumsporozoite protein, the obic membrane-anchoring sequence.

C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
 C:Keywords: sporozoite; surface antigen; tandem repeat
 F:1-19/Domain: signal sequence #status predicted <SIG>
 F:20-419/Product: circumsporozoite protein #status predicted <MAT>
 F:99-310/Region: 4-residue repeats
 F:344-397/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match 79.6%; Score 82; DB 1; Length 419;
 Best Local Similarity 72.2%; Pred. No. 0.00012;
 Matches 13; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EWSPCSVTGNGIQVRIK 18
 |||||
 Db 353 EWSPCSVTGKGVRRRK 370

RESULT 15

T21371
 hypothetical protein F25H8.3 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999

C/Accession: T21371; T24896
 R:Gajadaty, S.
 submitted to the EMBL Data Library, February 1996
 A:Reference number: Z19413

A:Accession: T21371
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-2165 <WIL>
 A:Cross-references: EMBL:Z69360; PIDN:CAA93287.1; GSPDB:GN00022; CESP:F25H8.3
 A:Experimental source: clone F25H8

R:Gajadaty, S.
 submitted to the EMBL Data Library, February 1996
 A:Reference number: Z19949

A:Accession: T24896
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-2165 <WIL>
 A:Cross-references: EMBL:Z69361; PIDN:CAA93288.1; GSPDB:GN00022; CESP:F25H8.3
 A:Experimental source: clone T13H10

C:Genetics:
 A:Gene: CESP:F25H8.3
 A:Map position: 4
 A:Introns: 31/1; 52/1; 135/2; 193/3; 216/1; 266/1; 495/2; 547/3; 584/3; 634/2; 744/1; 81/
 Query Match 77.7%; Score 80; DB 2; Length 2165;

Best Local Similarity 86.7%; Pred. No. 0.00097;
Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 WSPCSVTGNGIQR 16
|||||:|||||
Db 1304 WSPCSVTGSGIQR 1318

Search completed: December 23, 2003, 15:34:25
Job time : 21.25 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: December 23, 2003, 15:33:08 ; Search time 15 Seconds

(without alignments)
56.432 Million cell updates/sec

Title: US-09-980-564-13

Perfect score: 103
Sequence: 1 EWSPCSVTGNGIGVRIK 18

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database: SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	103	100.0	388	1 CSP_PLARB	P26694 plasmodium
2	103	100.0	397	1 CSP_PLAFO	P19597 plasmodium
3	103	100.0	412	1 CSP_PLAFA	P02893 plasmodium
4	103	100.0	424	1 CSP_PLAFT	P13814 plasmodium
5	103	100.0	442	1 CSP_PLAFW	P08307 plasmodium
6	87	84.5	351	1 CSP_PLAKU	P04922 plasmodium
7	87	84.5	363	1 CSP_PLAKH	P02894 plasmodium
8	83	80.6	429	1 CSP_PLABR	P14593 plasmodium
9	83	80.6	429	1 CSP_PLAMA	P13815 plasmodium
10	82	79.6	378	1 CSP_PLACB	P06672 plasmodium
11	82	79.6	398	1 CSP_PLACC	P06673 plasmodium
12	82	79.6	419	1 CSP_PLACM	P06676 plasmodium
13	79	76.7	378	1 CSP_PLACL	P06675 plasmodium
14	79	76.7	401	1 CSP_PLACG	P06674 plasmodium
15	77	74.8	343	1 CSP_PLAVS	P13826 plasmodium
16	77	74.8	378	1 CSP_PLAVB	P06677 plasmodium
17	77	74.8	386	1 CSP_PLAVI	P06678 plasmodium
18	77	74.8	1170	1 TSP1_BOVIN	Q03110 plasmodium
19	77	74.8	1173	1 TSP1_XENLA	Q28178 bos taurus
20	76	73.8	367	1 CSP_PLAYO	P06914 plasmodium
21	74	71.8	339	1 CSP_PLABE	P06915 plasmodium
22	74	71.8	347	1 CSP_PLABA	P23093 plasmodium
23	73	70.9	559	1 TRAP_PLAFA	P16893 plasmodium
24	73	70.9	1170	1 TSP1_MOUSE	P35441 mus musculus
25	72.5	70.4	1572	1 BA12_HUMAN	Q060241 homo sapien
26	72	69.9	1074	1 SMSA_HUMAN	Q13591 homo sapien
27	72	69.9	1077	1 SMSA_MOUSE	Q62217 mus musculus
28	72	69.9	1522	1 BA13_HUMAN	Q06242 homo sapien
29	70	68.0	803	1 FSPD_XENLA	P35447 xenopus lae
30	70	68.0	1205	1 AT52_BOVIN	P79331 b adams-2
31	70	68.0	1207	1 AT19_HUMAN	Q8c59 homo sapien
32	70	68.0	1210	1 AT19_MOUSE	P59509 mus musculus
33	70	68.0	1211	1 AT52_HUMAN	O95450 h adams-2

34	69	67.0	807	1 FSPD_RAT	P35446 rattus norv
35	69	67.0	1170	1 TSP1_HUMAN	P07996 homo sapien
36	68	66.0	562	1 AT15_MOUSE	P59384 mus musculus
37	67	65.0	437	1 PROP_MOUSE	P11680 mus musculus
38	67	65.0	1172	1 TSP2_HUMAN	P35442 homo sapien
39	67	65.0	1172	1 TSP2_MOUSE	Q03350 mus musculus
40	67	65.0	1178	1 TSP2_CHICK	P35440 gallus gall
41	67	65.0	1205	1 AT53_HUMAN	O15072 homo sapien
42	67	65.0	1911	1 AT20_HUMAN	P59510 homo sapien
43	66	64.1	347	1 CTGF_RAT	O9x169 rattus norv
44	66	64.1	348	1 CTGF_MOUSE	P23268 mus musculus
45	66	64.1	349	1 CTGF_BOVIN	O18739 bos taurus

ALIGNMENTS

```

RESULT 1
CSP_PLARB          STANDARD;          PRT;          388 AA.
ID CSP_PLARB
AC P26694:
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Circumsporozoite protein precursor (CS).
OS Plasmodium reichenowi.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_Taxid=5864;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=91201303; Pubmed=2016283;
RA Lal A.A., Goldman I.F.;
RT "Circumsporozoite protein gene from Plasmodium reichenowi, a
RT chimpanzee malaria parasite evolutionarily related to the human
RT malaria parasite Plasmodium falciparum."
RL J. Biol. Chem. 266:6686-6689(1991).
CC -!- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
CC SURFACE ANTIGEN ON THE SPOOROZOITE (THE INFECTIVE STAGE OF THE
CC MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
CC VERTEBRATE HOST).
CC -!- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
CC ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
CC WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -!- SIMILARITY: Contains 1 TSP type-1 domain.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; M60972; AAA29561.1; -.
DR PIR; A39756; A39756.
DR InterPro; IPR001067; Circumsporozoite.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00090; TSP_1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PS50092; TSP1; 1.
KW Malaria; Sporozoite; Repeat; Signal.
FT SIGNAL 1..16
FT CHAIN 17..388
FT DOMAIN 120..267
FT DOMAIN 313..366
FT DOMAIN 62 X 4 AA TANDEM REPEATS OF N-A-N-P.
SQ SEQUENCE 388 AA; 42245 MW; C031BFB2E35604 CRC64;

Query Match          100.0%; Score 103; DB 1; Length 388;
Best Local Similarity 100.0%; Pred. No. 1.3e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 EWSPCSVTGNGIGVRIK 18

```

Db 321 EMSPCSVTCGNGIQRK 338

RESULT 2

CSP_PLAFO STANDARD; PRT; 397 AA.
AC P19597; Q25798;
DT 01-FEB-1991 (Rel. 17, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Circumsporozoite protein precursor (CS).
OS Plasmodium falciparum (isolate NF54).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
NCBI_TaxID=5843;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89345189; PubMed=2668995;
RA Campbell J.R.;
RT "DNA sequence of the gene encoding a Plasmodium falciparum malaria candidate vaccine antigen";
RL Nucleic Acids Res. 17:5854-5854 (1989).
RN [2]
RP REVISIONS.
RA Campbell J.R.;
RL Submitted (JUL-1997) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=92155298; PubMed=1346766;
RA Davis J.R., Cortese J.F., Herrington D.A., Murphy J.R., Clyde D.F., Thomas A.W., Bagar S., Cochran M.A., Thanassi J., Levine M.M., Hackett C.S.;
RT "Plasmodium falciparum: in vitro characterization and human infectivity of a cloned line";
RL Exp. Parasitol. 74:159-168 (1992).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=89364998; PubMed=2671723;
RA Capere P., Gentz R., Matile H., Pink J.R., Sinigaglia F.;
RT "The circumsporozoite protein gene from NF54, a Plasmodium falciparum isolate used in malaria vaccine trials";
RL Mol. Biochem. Parasitol. 35:185-190 (1989).
CC -1- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT SURFACE ANTIGEN ON THE SPOOROZOITE (THE INFECTIVE STAGE OF THE MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE VERTEBRATE HOST).
CC -1- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -1- SIMILARITY: Contains 1 TSP type-1 domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X15363; CAA33421.1; -
CC EMBL; M83886; AAA29521.1; -
CC EMBL; M22982; AAA29527.1; -
CC PIR; S05428; S05428.
CC InterPro; IPR003067; Circmsprzoite.
CC InterPro; IPR000884; TSP1.
CC Pfam; PF00090; tsp_1; 1.
CC PRINTS; PR01303; CRCMSPRZOITE.
CC SMART; SMO0209; TSP1.1.
CC PROSITE; PS50092; TSP1.1.
CC Malaria; Sporozoite; Repeat; Signal.
CC SIGNAL 1 16 PROBABLE.
CC CHAIN 17 397 CIRCUMSPOROZOITE PROTEIN.
CC DOMAIN 101 272 43 X 4 AA TANDEM REPEATS OF N-A-N-P.

FT DOMAIN 322 375 TSP TYPE-1.
FT CONFLICT 194 A -> ANPNANPNA (IN REF. 4).
SQ SEQUENCE 397 AA; 42646 MW; 9E81146F59BCEA3 CRC64;

Query Match 100.0%; Score 103; DB 1; Length 397;
Best Local Similarity 100.0%; Pred. No. 1.3e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EMSPCSVTCGNGIQRK 18
Db 330 EMSPCSVTCGNGIQRK 347

RESULT 3

CSP_PLAFA STANDARD; PRT; 412 AA.
AC P02893;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Circumsporozoite protein precursor (CS).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84250215; PubMed=6204383;
RA Dame J.B., Williams J.L., McCutchan T.F., Weber J.L., Wirtz R.A., Hochmeyer W.T., Maloy W.L., Haynes J.D., Schneider I., Roberts D., Sanders G.S., Reddy E.P., Diggs C.L., Miller L.H.;
RT "Structure of the gene encoding the immunodominant surface antigen on the sporozoite of the human malaria parasite Plasmodium falciparum";
RL Science 225:593-599 (1984).
CC -1- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT SURFACE ANTIGEN ON THE SPOOROZOITE (THE INFECTIVE STAGE OF THE MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE VERTEBRATE HOST).
CC -1- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -1- SIMILARITY: Contains 1 TSP type-1 domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
CC EMBL; K02194; AAA29524.1; -
CC PIR; A03388; OZZOAF.
CC InterPro; IPR003067; Circmsprzoite.
CC InterPro; IPR000884; TSP1.
CC Pfam; PF00090; tsp_1; 1.
CC PRINTS; PR01303; CRCMSPRZOITE.
CC SMART; SMO0209; TSP1.1.
CC PROSITE; PS50092; TSP1.1.
CC Malaria; Sporozoite; Repeat; Signal.
CC SIGNAL 1 16 PROBABLE.
CC CHAIN 17 412 CIRCUMSPOROZOITE PROTEIN.
CC DOMAIN 123 290 41 X 4 AA TANDEM REPEATS OF P-N-A-N.
CC FT 337 390 TSP TYPE-1.
CC SQ SEQUENCE 412 AA; 44420 MW; 1EEED3DE90965F8 CRC64;
Query Match 100.0%; Score 103; DB 1; Length 412;
Best Local Similarity 100.0%; Pred. No. 1.4e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EMSPCSVTCGNGIQRK 18
Db 345 EMSPCSVTCGNGIQRK 362

```

RESULT 4
CSP_PLAFT ID CSP_PLAFT STANDARD: PRT; 424 AA.
AC P13814:
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Circumsporozoite protein precursor (CS).
OS Plasmodium falciparum (isolate t4 / Thailand).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5846;
RN [1]
RP MEDLINE=87115205; PubMed=3306373;
RA del Portillo H.A., Nussenzweig R.S., Bnea V.;
RT "Circumsporozoite gene of a Plasmodium falciparum strain from Thailand."
RL Mol. Biochem. Parasitol. 24:289-294(1987).
CC -1- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT SURFACE ANTIGEN ON THE SPOROZOITE (THE INFECTIVE STAGE OF THE MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE VERTEBRATE HOST).
CC -1- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -1- SIMILARITY: Contains 1 TSP type-1 domain.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch).
-----
CC EMBL; M19752; AAA29555.1; -.
DR PIR; A54533; A54533.
DR InterPro; IPR003067; Circmsprzoite.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PS50092; TSP1; 1.
DR Malaria; Sporozoite; Repeat; Signal.
FT SIGNAL 1 16 PROBABLE.
FT CHAIN 17 424 CIRCUMSPOROZOITE PROTEIN.
FT DOMAIN 123 300 45 X 4 AA TANDEM REPEATS OF N-A-N-P.
FT DOMAIN 349 402 TSP TYPE-1.
SQ SEQUENCE 424 AA; 45610 MW; 710AB14238786CD9 CRC64;

Query Match 100.0%; Score 103; DB 1; Length 424;
Best Local Similarity 100.0%; Pred. No. 1.4e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EMSPCSVTCGNGIOYRIK 18
DB 357 EMSPCSVTCGNGIOYRIK 374

```

```

RX MEDLINE=87115616; PubMed=3543671;
RA Lockyer M.J., Schwarz R.T.;
RT "Strain variation in the circumsporozoite protein gene of Plasmodium falciparum."
RL Mol. Biochem. Parasitol. 22:101-108(1987).
CC -1- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT SURFACE ANTIGEN ON THE SPOROZOITE (THE INFECTIVE STAGE OF THE MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE VERTEBRATE HOST).
CC -1- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -1- SIMILARITY: Contains 1 TSP type-1 domain.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch).
-----
CC EMBL; M15505; AAA29554.1; -.
DR PIR; A54529; A54529.
DR InterPro; IPR003067; Circmsprzoite.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PS50092; TSP1; 1.
DR Malaria; Sporozoite; Repeat; Signal.
FT SIGNAL 1 16
FT CHAIN 17 442 CIRCUMSPOROZOITE PROTEIN
FT DOMAIN 130 320 47 X 4 AA TANDEM REPEATS OF N-A-N-P.
FT DOMAIN 367 420 TSP TYPE-1.
SQ SEQUENCE 442 AA; 47402 MW; BD57A9A152B85E03 CRC64;

Query Match 100.0%; Score 103; DB 1; Length 442;
Best Local Similarity 100.0%; Pred. No. 1.4e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EMSPCSVTCGNGIOYRIK 18
DB 375 EMSPCSVTCGNGIOYRIK 392

RESULT 6
CSP_PLAFT ID CSP_PLAFT STANDARD: PRT; 351 AA.
AC P04522:
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Circumsporozoite protein precursor (CS).
OS Plasmodium knowlesi (strain nur1).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5652;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85272582; PubMed=4023712;
RA Sharma S., Svec P., Mitchell G.H., Godson G.N.;
RT "Diversity of circumsporozoite antigen genes from two strains of the malarial parasite Plasmodium knowlesi."
RL Science 229:779-782(1985).
CC -1- MISCELLANEOUS: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT SURFACE ANTIGEN ON THE SPOROZOITE (THE INFECTIVE STAGE OF THE MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE VERTEBRATE HOST).
CC -1- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -1- SIMILARITY: Contains 1 TSP type-1 domain.
-----

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M11031; AAA29540.1; -.
DR PIR; A26253; OZZOAK.
DR InterPro; IPR003067; Circmeprozote.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PS50092; TSP1; 1.
KW Malaria; Sporozote; Repeat; Signal.
FT SIGNAL 1 19 PROBABLE.
FT CHAIN 20 351 CIRCUMSPOROZOITE PROTEIN.
FT DOMAIN 101 235 15 X 9 AA TANDEM REPEATS OF A-A-G-A-G-G-
E-O-P.
FT DOMAIN 277 329 TSP TYPE-1.
SQ SEQUENCE 351 AA; 34782 MW; A85E87A152E6485B CRC64;

Query Match 84.5%; Score 87; DB 1; Length 351;
Best Local Similarity 72.2%; Pred. No. 3.4e-06;
Matches 13; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 EWSPCSVTCGNGVIRIK 18
ID CSP_PLAKH STANDARD; PRT; 363 AA.
AC P02894;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Circumsporozoite protein precursor (CS).
OS Plasmodium knowlesi (strain H).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5651;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84026486; Pubmed=6313209;
RA Ozaki L.S., Svec P., Nussenzweig V., Godson G.N.;
RT "Structure of the plasmodium knowlesi gene coding for the
RT circumsporozoite protein."
RL Cell 34:815-822(1983).
RN [2]
RN SEQUENCE OF 84-258 FROM N.A.
RX MEDLINE=83297689; Pubmed=6193427;
RA Godson G.N., Ellis J., Svec P., Schlesinger D.H., Nussenzweig V.;
RT "Identification and chemical synthesis of a tandemly repeated
RT immunogenic region of Plasmodium knowlesi circumsporozoite protein."
RL Nature 305:29-33(1983).
CC -1- MISCELLANEOUS: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
CC SURFACE ANTIGEN ON THE SPOROZOITE (THE INFECTIVE STAGE OF THE
CC MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
CC VERTEBRATE HOST).
CC -1- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
CC ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
CC WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -1- SIMILARITY: Contains 1 TSP type-1 domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

```

```

CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; K00823; AAA19699.1; -.
DR EMBL; K00772; AAA29556.1; -.
DR PIR; A90841; OZZOAK.
DR InterPro; IPR003067; Circmeprozote.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PS50092; TSP1; 1.
KW Malaria; Sporozote; Repeat; Signal.
FT SIGNAL 1 19 PROBABLE.
FT CHAIN 20 363 CIRCUMSPOROZOITE PROTEIN.
FT DOMAIN 98 241 12 X 12 AA TANDEM REPEATS OF N-A-G-O-P-Q-
A-O-G-D-G-A.
FT DOMAIN 289 341 TSP TYPE-1.
SQ SEQUENCE 363 AA; 36793 MW; 574DF4BD320A7955 CRC64;

Query Match 84.5%; Score 87; DB 1; Length 363;
Best Local Similarity 72.2%; Pred. No. 3.5e-06;
Matches 13; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 EWSPCSVTCGNGVIRIK 18
ID CSP_PLABR STANDARD; PRT; 393 AA.
AC P14593;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Circumsporozoite protein (CS) (Fragment).
OS Plasmodium brasilianum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5624;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88186854; Pubmed=3128542;
RA Lal A.A., Ja Cruz V.F., Collins W.E., Campbell G.H., Procell P.M.,
RA McCutchan T.F.;
RT "Circumsporozoite protein gene from Plasmodium brasilianum. Animal
RT reservoirs for human malaria parasites?";
RL J. Biol. Chem. 263:5495-5498(1988).
CC -1- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
CC SURFACE ANTIGEN ON THE SPOROZOITE (THE INFECTIVE STAGE OF THE
CC MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
CC VERTEBRATE HOST).
CC -1- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
CC ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
CC WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -1- SIMILARITY: Contains 1 TSP type-1 domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J03203; AAA29553.1; -.
DR PIR; A60610; A60610.
DR InterPro; IPR003067; Circmeprozote.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PS50092; TSP1; 1.
KW Malaria; Repeat; Sporozote.

```

FT NON_TER 1 1
 FT DOMAIN 16 278 TANDEM REPEATS OF G-N-A-A.
 FT DOMAIN 319 371 TSP TYPE-1.
 SQ SEQUENCE 393 AA; 35372 MW; B32944419BC600AA CRC64;

Query Match 80.6%; Score 83; DB 1; Length 393;
 Best Local Similarity 77.8%; Pred. No. 1.5e-05;
 Matches 14; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 EMSPCSTVCNGIGVRIK 18
 |||||
 Db 327 EMSPCSTVCNGIGVRIK 344

RESULT 9

CSP_PLAMA STANDARD; PRT; 429 AA.

AC P13815;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Circumsporozoite protein precursor (CS).
 OS Plasmodium malariae.
 CC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 CC NCBI_TaxID=5858;
 CC [1]
 CC SEQUENCE FROM N.A.
 CC MEDLINE=89640027; PubMed=3054537;
 CC Lai A.A., Ja Cruz V.F., Campbell G.H., Procell P.M., Collins W.E.,
 CC McCutchan T.F.;
 CC "Structure of the circumsporozoite gene of Plasmodium malariae.";

RT Mol. Biochem. Parasitol. 30:291-294(1988).
 CC -1- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
 CC SURFACE ANTIGEN ON THE SPOROZOITE (THE INFECTIVE STAGE OF THE
 CC MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
 CC VERTEBRATE HOST).

CC -1- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
 CC ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
 CC WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
 CC -1- SIMILARITY: Contains 1 TSP type-1 domain.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: J03992; AAA29557.1; -;
 CC PIR: A54504; A54504.
 CC InterPro: IPR003067; Crcmpzrzoite.
 CC InterPro: IPR000884; TSP1.
 CC Pfam: PF00090; TSP_1; 1.
 CC PRINTS: PRO1303; CIRCMSPRZOITE.
 CC SMART: SM00209; TSP1.1.
 CC PROSITE: PS50092; TSP1.1.

DR Malaria; Sporozoite; Repeat; Signal.
 KW SIGNAL
 FT SIGNAL 1 15 PROBABLE.
 FT CHAIN 16 429 CIRCUMSPOROZOITE PROTEIN.
 FT DOMAIN 110 314 4 AA TANDEM REPEATS.
 FT DOMAIN 355 407 TSP TYPE-1.
 SQ SEQUENCE 429 AA; 41596 MW; 3629D641D1C0B7E CRC64;

Query Match 80.6%; Score 83; DB 1; Length 429;
 Best Local Similarity 77.8%; Pred. No. 1.7e-05;
 Matches 14; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 EMSPCSTVCNGIGVRIK 18
 |||||
 Db 363 EMSPCSTVCNGIGVRIK 380

RESULT 10
 CSP_PLACC STANDARD; PRT; 378 AA.

AC P08672;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Circumsporozoite protein precursor (CS).
 OS Plasmodium cynomolgi (strain Berck).
 CC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 CC NCBI_TaxID=5828;
 CC [1]
 CC SEQUENCE FROM N.A.

CC MEDLINE=87102878; PubMed=3802196;
 CC Galinski M.R., Arnot D.E., Cochrane A.H., Barnwell J.W.,
 CC Nussenzweig R.S., Enea V.,
 CC "The circumsporozoite gene of the Plasmodium cynomolgi complex";
 CC Cell 48:311-319(1987).
 CC -1- DOMAIN: THERE ARE 10 TANDEM COPIES OF A 9-RESIDUE REPEAT (PRECEDED
 CC BY A 6-RESIDUE INCOMPLETE REPEAT) AND 3 TANDEM COPIES OF A 16-
 CC RESIDUE REPEAT (FOLLOWED BY 3 SHORTER, INCOMPLETE COPIES).

CC -1- MISCELLANEOUS: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
 CC SURFACE ANTIGEN ON THE SPOROZOITE (THE INFECTIVE STAGE OF THE
 CC MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
 CC VERTEBRATE HOST).
 CC -1- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
 CC ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
 CC WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
 CC -1- SIMILARITY: Contains 1 TSP type-1 domain.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: M15104; AAA29532.1; -;
 CC PIR: D26255; OZQAB.
 CC InterPro: IPR003067; Crcmpzrzoite.
 CC InterPro: IPR000884; TSP1.
 CC Pfam: PF00090; TSP_1; 1.
 CC PRINTS: PRO1303; CIRCMSPRZOITE.
 CC SMART: SM00209; TSP1.1.

DR PROSITE: PS50092; TSP1.1.
 KW Malaria; Sporozoite; Repeat; Signal.
 FT SIGNAL 1 19 PROBABLE.
 FT CHAIN 20 378 CIRCUMSPOROZOITE PROTEIN.
 FT DOMAIN 97 192 10.5 X 9 AA REPEATS.
 FT REPEAT 193 240 3 X 16 AA TANDEM REPEATS.
 FT REPEAT 241 251
 FT REPEAT 252 260
 FT REPEAT 261 268
 FT DOMAIN 304 356

SQ SEQUENCE 378 AA; 36286 MW; 779BA081C140793F CRC64;
 TSP TYPE-1.

Query Match 79.6%; Score 82; DB 1; Length 378;
 Best Local Similarity 72.2%; Pred. No. 2.1e-05;
 Matches 13; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 EMSPCSTVCNGIGVRIK 18
 |||||
 Db 312 EMSPCSTVCNGIGVRIK 329

RESULT 11

CSP_PLACC STANDARD; PRT; 398 AA.

AC P08673;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)

```

DE Circumsporozoite protein precursor (CS).
OS Plasmodium cynomolgi (strain Ceylon).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5829;
RN [1]
RP MEDLINE=67102878; PubMed=3802196;
RA Galinski M.R., Arnot D.E., Cochrane A.H., Barnwell J.W.,
RT Nussenzweig R.S., Enea V.;
RL "The circumsporozoite gene of the Plasmodium cynomolgi complex.";
RL Cell 48:311-319(1987).
CC -1- MISCELLANEOUS: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
CC SURFACE ANTIGEN ON THE SPOOROZITE (THE INFECTIVE STAGE OF THE
CC MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
CC VERTEBRATE HOST).
CC -1- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
CC ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
CC WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -1- SIMILARITY: Contains 1 TSP type-1 domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M5103; AAA29533.1; -
DR InterPro; IPR003067; Circmepzroite.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; CIRCMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PS50092; TSP1; 1.
KW Malaria; Sporozoite; Repeat; Signal.
FT SIGNAL 19 PROBABLE.
FT CHAIN 20 398 CIRCUMSPOROZOITE PROTEIN.
FT DOMAIN 97 249 17 X 9 AA TANDEM REPEATS OF A-G-N-N-A-A-
FT DOMAIN 324 376 TSP TYPE-1.
FT SEQUENCE 324 376 A-G-E.
SQ SEQUENCE 398 AA; 3718 MW; 6DFA2E8A62ED05BF CRC64;

Query Match 79.6%; Score 82; DB 1; Length 398;
Best Local Similarity 72.2%; Pred. No. 2.2e-05;
Matches 13; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EMSPCSVTCGKGVRRK 18
| | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | |
Db 332 EMSPCSVTCGKGVRRK 349

RESULT 12
CSP_PLACM STANDARD; PRT; 419 AA.
AC P08676;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Circumsporozoite protein precursor (CS).
OS Plasmodium cynomolgi (strain Mulligan/NIH).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5832;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=67102878; PubMed=3802196;
RA Galinski M.R., Arnot D.E., Cochrane A.H., Barnwell J.W.,
RT Nussenzweig R.S., Enea V.;
RL "The circumsporozoite gene of the Plasmodium cynomolgi complex.";
RL Cell 48:311-319(1987).
CC -1- MISCELLANEOUS: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
CC SURFACE ANTIGEN ON THE SPOOROZITE (THE INFECTIVE STAGE OF THE
CC MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE

```

```

CC VERTEBRATE HOST).
CC -1- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
CC ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
CC WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -1- SIMILARITY: Contains 1 TSP type-1 domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M5102; AAA29539.1; -
DR InterPro; IPR003067; Circmepzroite.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; CIRCMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PS50092; TSP1; 1.
KW Malaria; Sporozoite; Repeat; Signal.
FT SIGNAL 19 PROBABLE.
FT CHAIN 20 419 CIRCUMSPOROZOITE PROTEIN.
FT DOMAIN 99 314 54 X 4 AA TANDEM REPEATS OF N-A-[DG]-G.
FT DOMAIN 345 397 TSP TYPE-1.
SQ SEQUENCE 419 AA; 38924 MW; 8F46CD8A1B4BFF4 CRC64;

Query Match 79.6%; Score 82; DB 1; Length 419;
Best Local Similarity 72.2%; Pred. No. 2.3e-05;
Matches 13; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EMSPCSVTCGKGVRRK 18
| | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | |
Db 353 EMSPCSVTCGKGVRRK 370

RESULT 13
CSP_PLACM STANDARD; PRT; 378 AA.
AC P08675;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Circumsporozoite protein precursor (CS).
OS Plasmodium cynomolgi (strain London).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5831;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=67102878; PubMed=3802196;
RA Galinski M.R., Arnot D.E., Cochrane A.H., Barnwell J.W.,
RT Nussenzweig R.S., Enea V.;
RL "The circumsporozoite gene of the Plasmodium cynomolgi complex.";
RL Cell 48:311-319(1987).
CC -1- MISCELLANEOUS: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
CC SURFACE ANTIGEN ON THE SPOOROZITE (THE INFECTIVE STAGE OF THE
CC MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
CC VERTEBRATE HOST).
CC -1- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
CC ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
CC WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -1- SIMILARITY: Contains 1 TSP type-1 domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M5101; AAA29537.1; -

```

```

DR InterPro; IPR003067; Circmsprzoite.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PS50092; TSP1; 1.
KW Malaria; Sporozoite; Repeat; 1.
FT SIGNAL 1 19 PROBABLE.
FT CHAIN 20 378 CIRCUMSPOROZOITE PROTEIN.
FT DOMAIN 98 211 18 X 6 AA TANDEM REPEATS OF D-G-A-R-A-
[EA]
FT DOMAIN 212 277 6 X 11 AA TANDEM REPEATS OF G-N-Q-A-G-G-
Q-A-G-A-G.
FT DOMAIN 304 356 TSP TYPE-1.
SQ SEQUENCE 378 AA; 37462 MW; 8295A913C36420C5 CRC64;

Query Match
Best Local Similarity 76.7%; Score 79; DB 1; Length 378;
Matches 12; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 1 EWSPCSTVTCGNGIVRIK 18
DB 312 EWSPCSTVTCGNGIVRIK 329

RESULT 14
CSP_PLAVS STANDARD; PRT; 401 AA.
ID P08574;
AC 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Circumsporoite protein precursor (CS).
OS Plasmodium cynomolgi (strain Gombak).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OK NCBI_TaxID=5830;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87102878; PubMed=3802196;
RA Galinski M.R., Arnoc D.E., Cochran A.H., Barnwell J.W.,
RA Nuezenwey R.S., Enea V.;
RT "The circumsporoite gene of the Plasmodium cynomolgi complex.";
RL Cell 48:311-319(1987).
CC -1- MISCELLANEOUS: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
SURFACE ANTIGEN ON THE SPOOROZOITE (THE INFECTIVE STAGE OF THE
MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
VERTERATE HOST).
CC -1- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -1- SIMILARITY: Contains 1 TSP type-1 domain.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
CC EMBL; M15100; AAA29536.1; -.
DR InterPro; IPR003067; Circmsprzoite.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PS50092; TSP1; 1.
KW Malaria; Sporozoite; Repeat; 1.
FT SIGNAL 1 19 PROBABLE.
FT CHAIN 20 401 CIRCUMSPOROZOITE PROTEIN.
FT DOMAIN 98 278 17 X 11 AA TANDEM REPEATS OF [DG]-G-A-A-
A-A-G-G-G-N.
FT DOMAIN 327 379 TSP TYPE-1.

```

```

SQ SEQUENCE 401 AA; 36664 MW; 57D666268238503E CRC64;

Query Match
Best Local Similarity 76.7%; Score 79; DB 1; Length 401;
Matches 12; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 1 EWSPCSTVTCGNGIVRIK 18
DB 335 EWSPCSTVTCGNGIVRIK 352

RESULT 15
CSP_PLAVS STANDARD; PRT; 343 AA.
ID P13826;
AC 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Circumsporoite protein (CS) (Fragment).
OS Plasmodium vivax (strain Salvador I).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OK NCBI_TaxID=126793;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86070222; PubMed=2416057;
RA McCutchan T.F., Lal A.A., de la Cruz V.F., Miller L.H., Maloy W.L.,
RA Charoenvit V., Beaudoin R.L., Querriy P., Wistar R. Jr., Hoffman S.L.,
RA Hockmeyer W.T., Collins W.E., Wirth D.;
RT "Sequence of the immunodominant epitope for the surface protein on
sporozoites of Plasmodium vivax."
RL Science 230:1381-1383(1985).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=87194878; PubMed=2437120;
RA de la Cruz V.F., Lal A.A., Welsh J.A., McCutchan T.F.;
RT "Evolution of the immunodominant domain of the circumsporoite
protein gene from Plasmodium vivax. Implications for vaccines.";
RL J. Biol. Chem. 262:6464-6467(1987).
CC -1- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
SURFACE ANTIGEN ON THE SPOOROZOITE (THE INFECTIVE STAGE OF THE
MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
VERTERATE HOST).
CC -1- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -1- SIMILARITY: Contains 1 TSP type-1 domain.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
CC EMBL; M15100; AAA29536.1; -.
DR InterPro; IPR003067; Circmsprzoite.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PS50092; TSP1; 1.
KW Sporozoite; Malaria; Repeat.
FT NON TER 1 1
FT DOMAIN 63 243 19 X 9 AA TANDEM REPEATS OF P-G-D-R-A-D-
G-O-P.
FT DOMAIN 269 321 TSP TYPE-1.
SQ SEQUENCE 343 AA; 34155 MW; 308FDB5BC15DFC3 CRC64;

Query Match
Best Local Similarity 74.8%; Score 77; DB 1; Length 343;
Matches 12; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 EWSPCSTVTCGNGIVRIK 16
DB 277 EWSPCSTVTCGNGIVRIK 292

```

Search completed: December 23, 2003, 15:33:45
 Job time : 16 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 23, 2003, 15:33:08 ; Search time 42 Seconds

(without alignments)
110.594 Million cell updates/sec

Title: US-09-980-564-13

Perfect score: 103

Sequence: 1 EMSPCSVTCGNGICVRIK 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL_23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriaph:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	103	100.0	69	5	Q8MZK7 plasmodium
2	103	100.0	79	5	Q9U0Q2 plasmodium
3	103	100.0	80	5	Q9U0P3 plasmodium
4	103	100.0	80	5	Q9U0M1 plasmodium
5	103	100.0	80	5	Q9U0P7 plasmodium
6	103	100.0	80	5	Q9U0P9 plasmodium
7	103	100.0	80	5	Q9U0P1 plasmodium
8	103	100.0	80	5	Q9U0Q3 plasmodium
9	103	100.0	80	5	Q9U0Q0 plasmodium
10	103	100.0	80	5	Q9U0P9 plasmodium
11	103	100.0	80	5	Q9U0P9 plasmodium
12	103	100.0	80	5	Q9U0P5 plasmodium
13	103	100.0	80	5	Q9U0P5 plasmodium
14	103	100.0	80	5	Q9U0Q4 plasmodium
15	103	100.0	80	5	Q9U0Q4 plasmodium
16	103	100.0	80	5	Q9U0Q0 plasmodium

17	103	100.0	80	5	Q9U0P6	Q9U0P6 plasmodium
18	103	100.0	80	5	Q9U0P2	Q9U0P2 plasmodium
19	103	100.0	80	5	Q9U0P8	Q9U0P8 plasmodium
20	103	100.0	80	5	Q9U0P4	Q9U0P4 plasmodium
21	103	100.0	80	5	Q9U0V9	Q9U0V9 plasmodium
22	103	100.0	80	5	Q9U0W3	Q9U0W3 plasmodium
23	103	100.0	80	5	Q9U0Q1	Q9U0Q1 plasmodium
24	103	100.0	106	5	Q9U0P2	Q9U0P2 plasmodium
25	103	100.0	106	5	Q9U0P1	Q9U0P1 plasmodium
26	103	100.0	106	5	Q9U0P0	Q9U0P0 plasmodium
27	103	100.0	115	5	Q25837	Q25837 plasmodium
28	103	100.0	115	5	Q25839	Q25839 plasmodium
29	103	100.0	115	5	Q9U0P3	Q9U0P3 plasmodium
30	103	100.0	115	5	Q25835	Q25835 plasmodium
31	103	100.0	115	5	Q25836	Q25836 plasmodium
32	103	100.0	115	5	Q9U0W7	Q9U0W7 plasmodium
33	103	100.0	117	5	Q25795	Q25795 plasmodium
34	103	100.0	117	5	Q25796	Q25796 plasmodium
35	103	100.0	117	5	Q25797	Q25797 plasmodium
36	103	100.0	117	5	Q25794	Q25794 plasmodium
37	103	100.0	356	5	Q819G5	Q819G5 plasmodium
38	103	100.0	360	5	Q819H7	Q819H7 plasmodium
39	103	100.0	360	5	Q819H6	Q819H6 plasmodium
40	103	100.0	363	5	Q819H3	Q819H3 plasmodium
41	103	100.0	364	5	Q819G4	Q819G4 plasmodium
42	103	100.0	368	5	Q819G3	Q819G3 plasmodium
43	103	100.0	380	5	Q819G8	Q819G8 plasmodium
44	103	100.0	380	5	Q819G1	Q819G1 plasmodium
45	103	100.0	383	5	Q9GPN1	Q9GPN1 plasmodium

ALIGNMENTS

RESULT 1

ID	Q8MZK7	PRELIMINARY	PRT	69 AA
AC	Q8MZK7			
DT	01-OCT-2002 (TREMBLrel. 22, Created)			
DT	01-OCT-2002 (TREMBLrel. 22, Last sequence update)			
DT	01-MAR-2003 (TREMBLrel. 23, Last annotation update)			
DE	Circumsporozoite protein II (Fragment).			
GN	CSPII.			
OS	Plasmodium falciparum.			
OC	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.			
OX	NCBI_taxid=5833;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=FCCL/HN;			
RA	Chen H.H., Yu X.B., Mu Z.D., Xu J.;			
RT	"Gene cloning of circumsporozoite protein (CSP) II gene from			
RL	Plasmodium falciparum (FCCL/HN)."			
DR	Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AY093672; AAM19072.1; -			
DR	InterPro: IPR003067; Circmsporzoite.			
DR	InterPro: IPR000884; TSP1.			
DR	Pfam; PF00090; tep_1; 1.			
DR	PRINTS; PRO1303; CRCMSPRZOITE.			
DR	SMART; SM00209; TSP1, 1.			
DR	PROSITE; PS50092; TSP1, 1.			
FT	NON TER			
FT	1			
FT	1			
SQ	SEQUENCE 69 AA; 7670 MW; F3E1A54671FE6FF CRC64;			

Query Match 100.0%; Score 103; DB 5; Length 69;
Best Local Similarity 100.0%; Pred. No. 7.2e-10;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EMSPCSVTCGNGICVRIK 18
|||
Db 2 EMSPCSVTCGNGICVRIK 19

RESULT 2

```

Q9U0Q2          PRELIMINARY;      PRT;      79 AA.
ID   Q9U0Q2
AC   Q9U0Q2;
DT   01-MAY-2000 (TREMBLrel. 13, Created)
DT   01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT   01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE   Circumsporozoite protein (Fragment).
GN   CS.
OS   Plasmodium falciparum.
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=5833;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=D51;
RA   de Stricker K., Vuust J., Jepsen S., Oeuvray C., Theisen M.;
RT   "Sequence variation in the non-repeat region of the Plasmodium
RT   falciparum glutamate rich protein (GLURP) from Brazil, Senegalese, and
RT   Burmese field isolates and from laboratory strains.";
RL   Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR   EMBL; AJ269945; CAB64243.1; -.
DR   InterPro; IPR003067; CircmSprzoite.
DR   InterPro; IPR000884; TSP1.
DR   Pfam; PF00090; tsp_1; 1.
DR   PRINTS; PR01303; CRCMSPRZOITE.
DR   SMART; SM00209; TSP1; 1.
DR   PROSITE; PS50092; TSP1; 1.
FT   NON_TER      1
FT   NON_TER      79
SQ   SEQUENCE      79 AA; 8893 MW; 56FBFA76D859B416 CRC64;

Query Match          100.0%; Score 103; DB 5; Length 79;
Best Local Similarity 100.0%; Pred. No. 8.3e-10;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 EMSPCSVTCGNGIQVRK 18
Db      27 EMSPCSVTCGNGIQVRK 44

RESULT 3
Q9U0P3          PRELIMINARY;      PRT;      80 AA.
ID   Q9U0P3
AC   Q9U0P3;
DT   01-MAY-2000 (TREMBLrel. 13, Created)
DT   01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT   01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE   Circumsporozoite protein (Fragment).
GN   CS.
OS   Plasmodium falciparum.
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=5833;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=D405;
RA   de Stricker K., Vuust J., Jepsen S., Oeuvray C., Theisen M.;
RT   "Sequence variation in the non-repeat region of the Plasmodium
RT   falciparum glutamate rich protein (GLURP) from Brazil, Senegalese, and
RT   Burmese field isolates and from laboratory strains.";
RL   Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR   EMBL; AJ269969; CAB64188.1; -.
DR   InterPro; IPR003067; CircmSprzoite.
DR   InterPro; IPR000884; TSP1.
DR   Pfam; PF00090; tsp_1; 1.
DR   PRINTS; PR01303; CRCMSPRZOITE.
DR   SMART; SM00209; TSP1; 1.
DR   PROSITE; PS50092; TSP1; 1.
FT   NON_TER      1
FT   NON_TER      80
SQ   SEQUENCE      80 AA; 9033 MW; BA71EBE0DB03193B CRC64;

Query Match          100.0%; Score 103; DB 5; Length 80;
Best Local Similarity 100.0%; Pred. No. 8.4e-10;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 EMSPCSVTCGNGIQVRK 18
Db      27 EMSPCSVTCGNGIQVRK 44

RESULT 4
Q9TW01          PRELIMINARY;      PRT;      80 AA.
ID   Q9TW01
AC   Q9TW01;
DT   01-MAY-2000 (TREMBLrel. 13, Created)
DT   01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE   Circumsporozoite protein (Fragment).
GN   CS.
OS   Plasmodium falciparum.
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=5833;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=M11, and D7;
RA   de Stricker K., Vuust J., Jepsen S., Oeuvray C., Theisen M.;
RT   "Sequence variation in the non-repeat region of the Plasmodium
RT   falciparum glutamate rich protein (GLURP) from Brazil, Senegalese, and
RT   Burmese field isolates and from laboratory strains.";
RL   Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR   EMBL; AJ269960; CAB64240.1; -.
DR   InterPro; IPR003067; CircmSprzoite.
DR   InterPro; IPR000884; TSP1.
DR   Pfam; PF00090; tsp_1; 1.
DR   PRINTS; PR01303; CRCMSPRZOITE.
DR   SMART; SM00209; TSP1; 1.
DR   PROSITE; PS50092; TSP1; 1.
FT   NON_TER      1
FT   NON_TER      80
SQ   SEQUENCE      80 AA; 9123 MW; 4614EEC68F0B1434 CRC64;

Query Match          100.0%; Score 103; DB 5; Length 80;
Best Local Similarity 100.0%; Pred. No. 8.4e-10;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 EMSPCSVTCGNGIQVRK 18
Db      27 EMSPCSVTCGNGIQVRK 44

RESULT 5
Q9U0P7          PRELIMINARY;      PRT;      80 AA.
ID   Q9U0P7
AC   Q9U0P7;
DT   01-MAY-2000 (TREMBLrel. 13, Created)
DT   01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT   01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE   Circumsporozoite protein (Fragment).
GN   CS.
OS   Plasmodium falciparum.
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=5833;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=D4230;
RA   de Stricker K., Vuust J., Jepsen S., Oeuvray C., Theisen M.;
RT   "Sequence variation in the non-repeat region of the Plasmodium
RT   falciparum glutamate rich protein (GLURP) from Brazil, Senegalese, and
RT   Burmese field isolates and from laboratory strains.";
RL   Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR   EMBL; AJ269961; CAB64180.1; -.
DR   InterPro; IPR003067; CircmSprzoite.
DR   InterPro; IPR000884; TSP1.
DR   Pfam; PF00090; tsp_1; 1.
DR   PRINTS; PR01303; CRCMSPRZOITE.
DR   SMART; SM00209; TSP1; 1.

```

RT PROSITE, PSS0092; TSP1, 1.
 FT NON_TER 1
 FT NON_TER 80
 SQ SEQUENCE 80 AA; 9047 MW; BA769C90DB031C3E CRC64;

Query Match 100.0%; Score 103; DB 5; Length 80;
 Best Local Similarity 100.0%; Pred. No. 8.4e-10;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 EWSPCSTVTCGNGIOVRK 18
 |||||
 Db 27 EWSPCSTVTCGNGIOVRK 44

RESULT 6

ID Q9TW97 PRELIMINARY; PRT; 80 AA.

AC Q9TW97;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Circumsporozoite protein (Fragment).
 GN CS.

OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5833;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN-D4403, and DA259;
 RA de Stricker K., Vuust J., Jepsen S., Oeuvray C., Theisen M.;
 RT "Sequence variation in the non-repeat region of the Plasmodium
 RT falciparum glutamate rich protein (GLURP) from Brazil, Senegalese, and
 RT Burmese field isolates and from laboratory strains."
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL; AJ269968; CAB64187.1; -;
 DR EMBL; AJ269962; CAB64181.1; -;
 DR InterPro; IPR003067; CircmSprzoite.
 DR Pfam; PF00090; tsp_1; 1.
 DR PRINTS; PR01303; CRCMSPRZOITE.
 DR SMART; SMO0209; TSP1, 1.
 DR PROSITE; PSS0092; TSP1, 1.

FT NON_TER 1
 FT NON_TER 80
 SQ SEQUENCE 80 AA; 9000 MW; 036C7E319AFE13B3 CRC64;

Query Match 100.0%; Score 103; DB 5; Length 80;
 Best Local Similarity 100.0%; Pred. No. 8.4e-10;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 EWSPCSTVTCGNGIOVRK 18
 |||||
 Db 27 EWSPCSTVTCGNGIOVRK 44

RESULT 7

ID Q9U0P1 PRELIMINARY; PRT; 80 AA.

AC Q9U0P1;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Circumsporozoite protein (Fragment).
 GN CS.

OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5833;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN-B1896;
 RA MEDLINE=20542035, PubMed=11087922;
 RA de Stricker K., Vuust J., Jepsen S., Oeuvray C., Theisen M.;
 RT "Conservation and heterogeneity of the glutamate-rich protein (GLURP)

RT among field isolates and laboratory lines of Plasmodium falciparum.";
 RL Mol. Biochem. Parasitol. 111:123-130(2000).
 DR EMBL; AJ269978; CAB64197.1; -;
 DR InterPro; IPR003067; CircmSprzoite.
 DR InterPro; IPR000884; TSP1.
 DR Pfam; PF00090; tsp_1; 1.
 DR PRINTS; PR01303; CRCMSPRZOITE.
 DR SMART; SMO0209; TSP1, 1.
 DR PROSITE; PSS0092; TSP1, 1.

FT NON_TER 1
 FT NON_TER 80
 SQ SEQUENCE 80 AA; 9102 MW; BFC6C970CEFOFA3E CRC64;

Query Match 100.0%; Score 103; DB 5; Length 80;
 Best Local Similarity 100.0%; Pred. No. 8.4e-10;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 EWSPCSTVTCGNGIOVRK 18
 |||||
 Db 27 EWSPCSTVTCGNGIOVRK 44

RESULT 8

ID Q9U0Q3 PRELIMINARY; PRT; 80 AA.

AC Q9U0Q3;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Circumsporozoite protein (Fragment).
 GN CS.

OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5833;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=D25;
 RA de Stricker K., Vuust J., Jepsen S., Oeuvray C., Theisen M.;
 RT "Sequence variation in the non-repeat region of the Plasmodium
 RT falciparum glutamate rich protein (GLURP) from Brazil, Senegalese, and
 RT Burmese field isolates and from laboratory strains."
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL; AJ269942; CAB64168.1; -;
 DR EMBL; AJ269942; CAB64168.1; -;
 DR InterPro; IPR003067; CircmSprzoite.
 DR InterPro; IPR000884; TSP1.
 DR Pfam; PF00090; tsp_1; 1.
 DR PRINTS; PR01303; CRCMSPRZOITE.
 DR SMART; SMO0209; TSP1, 1.

DR PROSITE; PSS0092; TSP1, 1.
 FT NON_TER 1
 FT NON_TER 80
 SQ SEQUENCE 80 AA; 9014 MW; 077C7E319AFE13B3 CRC64;

Query Match 100.0%; Score 103; DB 5; Length 80;
 Best Local Similarity 100.0%; Pred. No. 8.4e-10;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 EWSPCSTVTCGNGIOVRK 18
 |||||
 Db 27 EWSPCSTVTCGNGIOVRK 44

RESULT 9

ID Q9TVQ0 PRELIMINARY; PRT; 80 AA.

AC Q9TVQ0;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Circumsporozoite protein (Fragment).
 GN CS.

OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D4393, and D4372;
RA de Stricker K., Vuust J., Jepsen S., Oeuvray C., Theisen M.;
RT "Sequence variation in the non-repeat region of the Plasmodium
falciparum glutamate rich protein (GLURP) from Brazil, Senegalese, and
RT Burmese field isolates and from laboratory strains."
RL Submitted (SEP-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL; AJ269966; CAB64185.1; -
DR EMBL; AJ269964; CAB64183.1; -
DR InterPro; IPR003067; Crmpsprzoite.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PS50092; TSP1; 1.
FT NON_TER 1 1
FT NON_TER 80 80
SQ SEQUENCE 80 AA; 9059 MW; A756D1FFC41C1C21 CRC64;

Query Match 100.0%; Score 103; DB 5; Length 80;
Best Local Similarity 100.0%; Pred. No. 8.4e-10;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 EMSPCSVTCGNGIQVRIK 18
Db 27 EMSPCSVTCGNGIQVRIK 44

RESULT 10
Q9U0P9 PRELIMINARY; PRT; 80 AA.
AC Q9U0P9;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Circumsporozoite protein (Fragment).
GN CS.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LI;
RA de Stricker K., Vuust J., Jepsen S., Oeuvray C., Theisen M.;
RT "Sequence variation in the non-repeat region of the Plasmodium
falciparum glutamate rich protein (GLURP) from Brazil, Senegalese, and
RT Burmese field isolates and from laboratory strains."
RL Submitted (SEP-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL; AJ269951; CAB64176.1; -
DR InterPro; IPR003067; Crmpsprzoite.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PS50092; TSP1; 1.
FT NON_TER 1 1
FT NON_TER 80 80
SQ SEQUENCE 80 AA; 9032 MW; ADEDE6F0E266AD98E CRC64;

Query Match 100.0%; Score 103; DB 5; Length 80;
Best Local Similarity 100.0%; Pred. No. 8.4e-10;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 EMSPCSVTCGNGIQVRIK 18
Db 27 EMSPCSVTCGNGIQVRIK 44

RESULT 11
Q9TVP9 PRELIMINARY; PRT; 80 AA.
ID Q9TVP9

AC Q9TVP9;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Circumsporozoite protein (Fragment).
GN CS.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=M6, M7, and M1;
RA de Stricker K., Vuust J., Jepsen S., Oeuvray C., Theisen M.;
RT "Sequence variation in the non-repeat region of the Plasmodium
falciparum glutamate rich protein (GLURP) from Brazil, Senegalese, and
RT Burmese field isolates and from laboratory strains."
RL Submitted (SEP-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL; AJ269955; CAB64242.1; -
DR EMBL; AJ269958; CAB64238.1; -
DR EMBL; AJ269959; CAB64239.1; -
DR EMBL; AJ269956; CAB64241.1; -
DR InterPro; IPR003067; Crmpsprzoite.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PS50092; TSP1; 1.
FT NON_TER 1 1
FT NON_TER 80 80
SQ SEQUENCE 80 AA; 9074 MW; ABF40C90DB1C033E CRC64;

Query Match 100.0%; Score 103; DB 5; Length 80;
Best Local Similarity 100.0%; Pred. No. 8.4e-10;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 EMSPCSVTCGNGIQVRIK 18
Db 27 EMSPCSVTCGNGIQVRIK 44

RESULT 12
Q9U0Q0 PRELIMINARY; PRT; 80 AA.
AC Q9U0Q0;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Circumsporozoite protein (Fragment).
GN CS.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D28;
RA de Stricker K., Vuust J., Jepsen S., Oeuvray C., Theisen M.;
RT "Sequence variation in the non-repeat region of the Plasmodium
falciparum glutamate rich protein (GLURP) from Brazil, Senegalese, and
RT Burmese field isolates and from laboratory strains."
RL Submitted (SEP-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL; AJ269948; CAB64173.1; -
DR InterPro; IPR003067; Crmpsprzoite.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PS50092; TSP1; 1.
FT NON_TER 1 1
FT NON_TER 80 80
SQ SEQUENCE 80 AA; 9046 MW; BA7689D18F031C3E CRC64;

Query Match 100.0%; Score 103; DB 5; Length 80;
Best Local Similarity 100.0%; Pred. No. 8.4e-10;